

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:12:37 ; Search time 46 Seconds
(without alignments)
3861.668 Million cell updates/sec

Title: US-09-981-947B-2
Perfect score: 2903
Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWIQAAATTTATKKEHPL 563

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp Vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2463	84.8	511	4 Q8WV30	Q8WV30 homo sapien
2	2242	77.2	538	11 Q7TSP0	Q7TSP0 mus musculus
3	1713.5	59.0	431	11 Q7TSK9	Q7TSK9 mus musculus
4	1681	57.9	416	11 Q9J3J2	Q9J3J2 mus musculus
5	1119.5	38.6	501	6 Q863Y9	Q863Y9 equus caball
6	1112.5	38.3	501	11 Q8RIN7	Q8RIN7 mus musculus
7	1101.5	37.9	501	6 Q8WMN1	Q8WMN1 ovis aries
8	1006	34.7	199	4 Q96P00	Q96P00 homo sapien
9	1001.5	34.5	499	4 Q8WY3	Q8WY3 homo sapien
10	999	34.4	496	4 Q8WY5	Q8WY5 homo sapien
11	999	34.4	503	4 Q8WY4	Q8WY4 homo sapien
12	994	34.2	503	4 Q8WYF9	Q8WYF9 homo sapien
13	914.5	31.5	220	11 Q8BZR3	Q8BZR3 mus musculus
14	899	31.0	489	13 Q8WSW5	Q8WSW5 eptatretus
15	895	30.8	492	13 Q918M1	Q918M1 oncorhynchus
16	894	30.8	505	13 Q8JFQ9	Q8JFQ9 oncorhynchus

17	894.5	30.5	493	13 Q801M9	Q801M9 xenopus lae
18	878.5	30.3	553	5 Q8MXW2	Q8MXW2 halocynthia
19	878	30.2	503	13 Q9DFV2	Q9DFV2 salmo trutt
20	875	30.1	509	6 Q8WJ15	Q8WJ15 equus cabal
21	851	29.3	481	5 Q9V409	Q9V409 drosophila
22	847	29.2	478	13 Q918L9	Q918L9 cyprinus ca
23	839.5	28.9	493	11 Q8BLF7	Q8BLF7 mus musculu
24	829.5	28.6	497	4 Q8TDB9	Q8TDB9 homo sapien
25	828	28.5	520	4 Q8TDB8	Q8TDB8 homo sapien
26	785	27.0	533	13 Q7ZT03	Q7ZT03 ctenopharyn
27	770	26.5	396	11 Q8BTM5	Q8BTM5 mus musculu
28	710	24.5	482	13 Q98TW6	Q98TW6 oncorhynch
29	703	24.2	510	5 Q44827	Q44827 caenorhabdi
30	697	24.0	492	5 Q44826	Q44826 caenorhabdi
31	697	24.0	500	5 Q27079	Q27079 taenia soli
32	664	22.9	516	5 Q44616	Q44616 caenorhabdi
33	663.5	22.9	492	5 Q26580	Q26580 schistosoma
34	648	22.3	507	5 Q9V310	Q9V310 drosophila
35	646	22.3	521	5 Q26579	Q26579 schistosoma
36	600	20.7	600	5 Q95TE8	Q95TE8 drosophila
37	600	20.7	647	5 Q81R16	Q81R16 drosophila
38	582	20.0	505	5 Q26581	Q26581 schistosoma
39	579	19.9	510	5 Q27072	Q27072 taenia soli
40	557.5	19.2	491	5 Q9U622	Q9U622 drosophila
41	547.5	18.9	476	5 Q8T3V4	Q8T3V4 drosophila
42	542.5	18.7	476	5 Q9V309	Q9V309 drosophila
43	530.5	18.3	516	5 Q9V919	Q9V919 drosophila
44	529	18.2	586	5 Q8MLN1	Q8MLN1 drosophila
45	524	18.1	546	10 Q93Z41	Q93Z41 arabidopsis

ALIGNMENTS

RESULT 1

Q8WV30	PRELIMINARY;	PRT;	511 AA.
ID	Q8WV30		
AC	Q8WV30;		
DT	01-MAR-2002 (TREMELrel. 20, Created)		
DT	01-MAR-2002 (TREMELrel. 20, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	Similar to solute carrier family 2 (Facilitated glucose transporter), member 9.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Ovary;		
RA	Strausberg R.;		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.		
DR	EMBL; BC018897; AAH18897.1; -.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0005215; F:transporter activity; IEA.		
DR	GO; GO:0008643; P:carbohydrate transport; IEA.		
DR	InterPro; IPR007114; MFS.		
DR	InterPro; IPR005828; Sub transporter.		
DR	InterPro; IPR003663; Sugar transport.		
DR	InterPro; IPR005829; Sug transporter.		
DR	Pfam; PF00083; sugar_tr; 1.		
DR	PRINTS; PR00171; SUGTRNSPORT.		
DR	TIGRFAMS; TIGR00879; SP; 1.		
DR	PROSITE; PS00850; MFS; 1.		
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.		
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.		
KW	Sugar transport; Transmembrane; Transport.		
SQ	SEQUENCE 511 AA; 55737 MW; C4BCB7138BA0F01E CRC64;		

Query Match 84.8%; Score 2463; DB 4; Length 511;

QY 248 SLFPLDPRVYLLLEKHEARAVKAFQTLGKADVSQVEEVLAEASHVQSRIRLVLEL 307
Db 209 ----- 208
QY 308 LEAPYRVQVTVITVMACYQLCGLNAINFYNTSIFGKAGIPPAKIPYVTLSTGDIETLA 367
Db 209 -----IWFTNSIFGKAGIPQDKIPYITLSTGDIETLA 241
QY 368 AVFSGLVIEHLGRRPLLIGGFGMLGFLPGTLTITLTIQDHAPWVYLSIVGILATIASFC 427
Db 242 AIFSGLVIERLGRREPLLIGGFGMALFFGTLTATLTIQDAPWVYLSIVCILATIASFC 301
QY 428 SGPGGIPPTLTGEEFQOSORPAAFIAGVNVNLSNFAVGLLPPFFQKSLDTCYFLVFATI 487
Db 302 SGPGGIPPTLTGEEFQOSORPAAFIAGVNVNLSNFAVGLLPPFFQKSLDTCYFLVFATI 361
QY 488 CITGAIYLYFVLPETKNTYABISQAFSKRNKAYPPEEKIDSAVTDAPASGPTTNTA 546
Db 362 CIAGATYFVLPETKNTYABISQAFSKRNKAYPPEVKADSAVTEEKANSQ-TEPDSS 419
RESULT 4
Q99J2 PRELIMINARY; PRT; 416 AA.
AC Q99J2; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
DE SLC2A9.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=9796;
SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; BC006076; AA06076.1; -.
DR MGD; MGI:2152844; SLC2a9.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Hypothetical protein; Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 416 AA; 45096 MW; 0A72AAA65BF7AF2B CRC64;
Query Match 57.9%; Score 1681; DB 11; Length 416;
Best Local Similarity 67.7%; Pred. No. 6.5e-115;
Matches 337; Conservative 26; Mismatches 27; Indels 108; Gaps 2;
QY 49 KDWSCSLVSLVSLAGAFSSFLGYNLSVNVNAPTYIKAPYNESWERRHRPIDPDTLTL 108
Db 15 QKWSFLVVAALVGAFGSSFLGYNLSVNVNAPTYIKAFYNGTWRRHGQIDPDTLTL 74
QY 109 LWSVTYSIFAIGLVGTLTYVMIGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLI 168
Db 75 LWSVTYSIFAIGLVGTLTYVMIGKVLGRKHTLLANGFAISALLMACSLRAGTPEMLI 134
QY 169 VGRFIMGIDGVALSVLPYLSISPKRIGSLGQVTAIFCIGVTGQLLGLPELLGKE 228
Db 135 VGRFIMGIDGVALSVLPYLSISPKRIGSLGQVTAIFCIGVTGQLLGLPELLGGR- 193
QY 229 STWPLYFGVTVVPAVQLLSLPLDSPRVLLLEKNEARAVKAFQTLGKADVSQVEE 288

Db 194 ----- 193
QY 289 VLAESHVQSRIRLVLELRLAPYRVQVTVITVMACYQLCGLNAINFYNTSIFGKAGI 348
Db 194 -----IWFTNSIFGKAGI 207
QY 349 PPAKIPYVTLSTGDIETLAIVFSLGVLIEHLGRRPLLIGGFGMLGFLPGTLTITLTIQDHA 408
Db 208 PQDKIPYITLSTGDIETLAIVFSLGVLIEHLGRRPLLIGGFGMALFFGTLTATLTIQDQA 267
QY 409 PWYPLYSIVGILATIASFCSPGGIPPTLTGEEFQOSORPAAFIAGVNVNLSNFAVGLL 468
Db 268 PWYPLYSIVGILATIASFCSPGGIPPTLTGEEFQOSORPAAFIAGVNVNLSNFAVGLL 327
QY 469 FPFQKSLDTCYFLVFATITGAIYLYFVLPETKNTYABISQAFSKRNKAYPPEEKID 528
Db 328 FPFQKSLDTCYFLVFATITGAIYLYFVLPETKNTYABISQAFSKRNKAYPPEVKAD 387
QY 529 SAVTDAPASGPTTNTA 546
Db 388 SAMTEEKANSQ-TEPDSS 404
RESULT 5
Q863Y9 PRELIMINARY; PRT; 501 AA.
ID Q863Y9;
AC Q863Y9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE GLUT5 fructose transporter.
DE SLC2A5.
GN Equus caballus (Horse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]_TaxID=9796;
SEQUENCE FROM N.A.
RA Fernandez-Castano Merediz E., Dyer J., Salmon K.S.H.,
RA Shirazi-Beechey S.P.;
RT "Molecular characterisation of fructose transporter (GLUT5) in equine small intestine.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Fernandez-Castano Merediz E.;
RT "Fructose transport in the brush-border membrane of equine small intestine.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ555215; CAD87604.1; -.
DR GO; GO:0006020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006843; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 501 AA; 54824 MW; B3E80ADE06C16295 CRC64;
Query Match 38.6%; Score 1119.5; DB 6; Length 501;
Best Local Similarity 44.4%; Pred. No. 1e-73;
Matches 216; Conservative 101; Mismatches 162; Indels 7; Gaps 2;

EMBL; AJ315928; CAC86964.1; ..
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0005622; C:intracellular; IEA.
DR GO:0005351; F:sugar porter activity; IEA.
DR GO:0003700; F:transcription factor activity; IEA.
DR GO:0005215; F:transporter activity; IEA.
DR GO:0008643; P:carbohydrate transport; IEA.
DR GO:0008355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR000005; HTHRAAC.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar transp.
DR InterPro: IPR005829; Sug transporter.
DR Pfam: PF00083; sugar tr; 1.
DR PRINTS: PR00171; SUGTRNSPORT.
DR TIGRFAM: TIGR00879; SP; 1.
DR PROSITE: PS00041; HTH ARAC_FAMILY_1; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR Sugar transport; Transmembrane; Transport.
KW SEQUENCE 501 AA; 55546 MW; 23F0B43E944C9AD5 CRC64;
Query Match 37.9%; Score 1101.5; DB 6; Length 501;
Best Local Similarity 45.7%; Pred.No. 2.2e-72; Indels 3; Gaps 1;
Matches 217; Conservative 97; Mismatches 158; Indels 3; Gaps 1;
QY 56 LIVASLAGAGSFLYGYNLVVNAPTYKAFYNESWERRGRPIDPTLLWSVTYS 115
Db 17 IVLALIAFGSFGQYNNVATINSSEFMKQFYNTYDYRGEYNEFYLLWSVTYS 76
QY 116 IFAIGLVGTLYVMIGKVLGRKHTLLANGFAISAAALMACSLOQAFEMIVGRFIMG 175
Db 77 MFPPGGFLGSLMVGPLVNNLGRKGTLLFNNTFSIVPALLMGFSLAKSPFEMIIVRLVG 136
QY 176 IDGVALSVLPMYLSISPKIRGSLGQVTAIFICIGVTGQLLGPPELLGKSTWPLYF 235
Db 137 ICAGLSNVVPMYGLGSLAPKNWKGALGVVPQPIITIGILVAQIFGRSLANEEGPILL 196
QY 236 GVIVPAVQVQLLSLPPLPSPRYLLEKNEARAVKAFOTFLGKADVSQEVSEVLAESHV 295
Db 197 GLTGIPAVQLLPFLPPSPFRVLLIQKDEEAKRALRLRGWHVDVDAEIEILEEDRA 256
QY 296 QRSIRLVSVLELRAPYRQVIVTWACYQCLGNALWFYNTSIFGKAGIPPAKIPY 355
Db 257 EKANGFISVLKFKMSLRQVSIIVLMAGQLSGVNAIYYVADQIYLSAGVKEDDVQY 316
QY 356 VTLSTGGIETLAAVFSGLVIEHGRPLIGGLGFLMGLPFGTUTITLTLODHAPVPLYLS 415
Db 317 VTAGTCAVNVLTWCAIFVVELMGRFRLLLLGFSVCFACCVLTGALAQDVISWMPYVS 376
QY 416 IVCILALIASFCSPGPIPIITGEPFQOSQRAAFIIAGTWNLSNFAVGLLPFIQKS 475
Db 377 IACVISYVIGHALGSPGIPALLVTEIFLOSSRPAAYVAGTVHLSNFTVGLVFPPIQVG 436
QY 476 LDYCYFLVAFATIGTGAIVLYFVLPTKNTYAEISQAFSKRNK--AYPPEKI 527
Db 437 LGAYSFVFAVCFLTVTYVIFLIIPETKSKTFIEINQIFIKMKVFGVHPEKEEL 491
RESULT 8
Q96P00 PRELIMINARY; PRT; 199 AA.
AC Q96P00
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Solute carrier family 2 member 9 (Fragment).
GN SUC2A9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;

SEQUENCE FROM N.A.
RC TISSUE=Articular cartilage;
RA Neema G., Richardson S., Bell S., Carter S., Mobasheri A.;
RT "Molecular characterization and cloning of glucose transporters in human articular chondrocytes.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF421859; AAL16939.1; ..
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0005215; F:transporter activity; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR Pfam: PF00083; sugar tr; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR Transmembrane.
KW NON_TER 199 199
FT NON_TER 199 199
SQ SEQUENCE 199 AA; 21541 MW; 17C0770C9B36F338 CRC64;
Query Match 34.7%; Score 1006; DB 4; Length 199;
Best Local Similarity 100.0%; Pred.No. 6.9e-66; Indels 0; Gaps 0;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 IKAFYNESWERRGRPIDPTLLWSVTYSIFAIGLVGTLYVMIGKVLGRKHTLLAN 144
Db 1 IKAFYNESWERRGRPIDPTLLWSVTYSIFAIGLVGTLYVMIGKVLGRKHTLLAN 60
QY 145 NGFAISAAALMACSLOQAFEMIVGRFIMGIDGVALSVLPMYLSISPKIRGSLGQV 204
Db 61 NGFAISAAALMACSLOQAFEMIVGRFIMGIDGVALSVLPMYLSISPKIRGSLGQV 120
QY 205 TAITFICIGVFTGQLLGPPELLGKSTWPLYFGVIVPAVQVQLLSLPPLPSPRYLLEKH 264
Db 121 TAITFICIGVFTGQLLGPPELLGKSTWPLYFGVIVPAVQVQLLSLPPLPSPRYLLEKH 180
QY 265 NEARAVKAFOTFLGKADVS 283
Db 181 NEARAVKAFOTFLGKADVS 199
RESULT 9
Q8WYM3 PRELIMINARY; PRT; 499 AA.
AC Q8WYM3
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Facilitative glucose transporter GLUT11-c.
GN SLC2A11-c.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu N., Minoshima S., Sasaki T.;
RT "Molecular cloning of a novel member of the glucose transporter family.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AB067444; BAB83506.1; ..
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0005351; F:sugar porter activity; IEA.
DR GO:0005215; F:transporter activity; IEA.
DR GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar transp.

DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRfam; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transport; Transmembrane; Transport.
 KW SEQUENCE 503 AA; 54460 MW; 61131DAB689D288B CRC64;
 SQ

Query Match 34.4%; Score 999; DB 4; Length 503;
 Best Local Similarity 42.1%; Pred. No. 6.9e-65;
 Matches 195; Conservative 94; Mismatches 174; Indels 0; Gaps 0;

QY 56 LLVASLAGAFSSFLYGNLSVNNAPTYIKAFYNESWERRHGRPIDPDTLLMSVTVS 115
 DB 20 LLTICAAIGGTFQFGYNLSINAPTLHIQFTNETWQARTGEPLPDHLVLLMSLVLS 79
 QY 116 IFAIGGLVGTLLVIMIGKVLGRKHTLLANNFSAALLMACSLQAGAFEMILVGRFIMG 175
 DB 80 LYPGLGLFALLAGPLAITLGRKKSLLVNNIFVWSAAILFGFSRKAGSFEMIMLGRLLVG 139
 QY 176 IDGGVALSVLPYLSISPKIRSGLSQVTAIPICIGVFTGQLLGLPELLGKESTWYLYF 235
 DB 140 VNAGVSMNIQPMYLGESAPKELRGAVAMSSAIFTALGIVMGQVGLRELLGGPQAWPLLL 199
 QY 236 GVIVPAVQVLLSLPDPSPRYLLEKNEARAVKAFOTFLGKADVSQVEVEVLAESHV 295
 DB 200 ASCLVPGALQLASLPLPSPRYLLIDCGTDEACTAALRQLRGSGDLAGELEBEERAA 259
 QY 296 QRSIRLVSVLELLRAPYVWQVTVITMACYQCLGNALVFTNSIFGKAGIPPAKIPY 355
 DB 260 CQCRARRPWEFLQHRALRRQVTSVLVGSAMELCGNDVSAYAVSSVFRKAGVPAKIQY 319
 QY 356 VTLSGTGIIETAAVPSGLVIEHLGRRLPGLLGGFGLMGLFFGTITLTLDHAPVVPYLS 415
 DB 320 ALIIGTSCCELLTAVVSCVWIERVGRVLLIGVSLMTCWGSIFTVALCQSSFPWTLYLA 379
 QY 416 IVGILAIASFCSGPGGIPFLTGEFFQOSQORPAAFIAGTVNLSNFAVGLLFFPIQXS 475
 DB 380 MACIFAFILSFSGIPAGVGTILATELFDQWAPACVCGALMIMILVGLGFFPIMEA 439
 QY 476 LDTYCFVLFATICITGAILYFVLPETKNTVYAEISQAFSKRN 518
 DB 440 LSHFLYVPFLGVCVCGAIYTGFLPETKGTFOEISKELHRLN 482

RESULT 12
 Q8WXP9 PRELIMINARY; PRT; 503 AA.
 AC Q8WXP9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sugar transporter GLUT11 long form.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu X., Li W., Freeze H.H.;
 RT "Cloning and Characterization of GLUT11, a Novel Sugar Transporter
 with Alternative Splicing Forms";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DE EMBL: AF43201; AAL39061.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRfam; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transport; Transmembrane; Transport.
 KW SEQUENCE 503 AA; 54546 MW; 2C682EAA3D7BB889 CRC64;
 SQ

Query Match 34.2%; Score 994; DB 4; Length 503;
 Best Local Similarity 41.9%; Pred. No. 1.6e-64;
 Matches 194; Conservative 94; Mismatches 175; Indels 0; Gaps 0;

QY 56 LLVASLAGAFSSFLYGNLSVNNAPTYIKAFYNESWERRHGRPIDPDTLLMSVTVS 115
 DB 20 LLTICAAIGGTFQFGYNLSINAPTLHIQFTNETWQARTGEPLPDHLVLLMSLVLS 79
 QY 116 IFAIGGLVGTLLVIMIGKVLGRKHTLLANNFSAALLMACSLQAGAFEMILVGRFIMG 175
 DB 80 LYPGLGLFALLAGPLAITLGRKKSLLVNNIFVWSAAILFGFSRKAGSFEMIMLGRLLVG 139
 QY 176 IDGGVALSVLPYLSISPKIRSGLSQVTAIPICIGVFTGQLLGLPELLGKESTWYLYF 235
 DB 140 VNAGVSMNIQPMYLGESAPKELRGAVAMSSAIFTALGIVMGQVGLRELLGGPQAWPLLL 199
 QY 236 GVIVPAVQVLLSLPDPSPRYLLEKNEARAVKAFOTFLGKADVSQVEVEVLAESHV 295
 DB 200 ASCLVPGALQLASLPLPSPRYLLIDCGTDEACTAALRQLRGSGDLAGELEBEERAA 259
 QY 296 QRSIRLVSVLELLRAPYVWQVTVITMACYQCLGNALVFTNSIFGKAGIPPAKIPY 355
 DB 260 CQCRARRPWEFLQHRALRRQVTSVLVGSAMELCGNDVSAYAVSSVFRKAGVPAKIQY 319
 QY 356 VTLSGTGIIETAAVPSGLVIEHLGRRLPGLLGGFGLMGLFFGTITLTLDHAPVVPYLS 415
 DB 320 ALIIGTSCCELLTAVVSCVWIERVGRVLLIGVSLMTCWGSIFTVALCQSSFPWTLYLA 379
 QY 416 IVGILAIASFCSGPGGIPFLTGEFFQOSQORPAAFIAGTVNLSNFAVGLLFFPIQXS 475
 DB 380 MACIFAFILSFSGIPAGVGTILATELFDQWAPACVCGALMIMILVGLGFFPIMEA 439
 QY 476 LDTYCFVLFATICITGAILYFVLPETKNTVYAEISQAFSKRN 518
 DB 440 LSHFLYVPFLGVCVCGAIYTGFLPETKGTFOEISKELHRLN 482

RESULT 13
 Q8BZR3 PRELIMINARY; PRT; 220 AA.
 AC Q8BZR3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Solute carrier family 2 (Fragment).
 DE SLC2A9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22354683; PubMed=12466851;
 RX The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).

```

DR EMBL; AK033725; BAC28448.1; -.
DR MGD; MGI-2152844; Slc2a9.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
FT NON TR 1
SQ SEQUENCE 220 AA; 23861 MW; 033C140B030E86B9 CRC64;

Query Match 31.5%; Score 914.5; DB 11; Length 220;
Best Local Similarity 85.2%; Pred. No. 3.9e-59;
Matches 178; Conservative 13; Mismatches 17; Indels 1; Gaps 1;

Qy 338 YTNISFGKAGIPAKIPYVTLSTGGIETLAAVPSGLVIEHLGRRPLLITGGFGLMGLFFGT 397
Db 1 YTNISFGKAGIPQDKIPYVTLSTGGIETLAAVPSGLVIEHLGRRPLLITGGFGLMALFFGT 60

Qy 398 LTITLTQDHAPWPVYLSIVGILAIATSCFSGGGIPFFILTGEFFQOSORPAAFIAGTV 457
Db 61 LTATLTQDHAPWPVYLSIVCIATATSCFSGGGIPFFILTGEFFQOSERPAAFIAGTV 120

Qy 458 NWLSNFAVGLLPFPIQKSLDYCFLVFATICTGAIYLYFVLPETKNRTYAIISQAFSKR 517
Db 121 NWLSNFAVGLLPFPIQKSLDSYCFLVFATICTAGATFYFVLPETKNRTYAIISQAFAKR 180

Qy 518 NKAYPPEKIDSAVTDAPASPTTNTA 546
Db 181 NKAQPEVKADSAMTEKANSQ-TEPDSS 208

RESULT 14
Q8WSW5 PRELIMINARY; PRT; 489 AA.
AC Q8WSW5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glucose transporter.
GN GLUT.
OS Eukretetus stoutii (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptaretinae; Eptaretus.
OC NCBI TAXID=7765;
EN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Yao S.Y.M., Ng A.M.L., Baldwin S.A., Young J.D.;
RT "Molecular cloning and functional characterization of an ancient pre-vertebrate glucose transporter (GLUT) from the Pacific hagfish (Eptaretus stoutii).";
RT (Eptaretus stoutii).";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AY059413; AAL27090.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTRANSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

```


FT	DOMAIN	1	5	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	6	26	1. (POTENTIAL).
FT	DOMAIN	27	64	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	65	85	2. (POTENTIAL).
FT	DOMAIN	86	93	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	94	114	3. (POTENTIAL).
FT	DOMAIN	115	124	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	125	145	4. (POTENTIAL).
FT	DOMAIN	146	153	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	154	174	5. (POTENTIAL).
FT	DOMAIN	175	183	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	184	204	6. (POTENTIAL).
FT	DOMAIN	205	269	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	270	290	7. (POTENTIAL).
FT	DOMAIN	291	304	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	305	325	8. (POTENTIAL).
FT	DOMAIN	326	334	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	335	355	9. (POTENTIAL).
FT	DOMAIN	356	363	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	364	384	10. (POTENTIAL).
FT	DOMAIN	385	399	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	400	420	11. (POTENTIAL).
FT	DOMAIN	421	424	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	425	445	12. (POTENTIAL).
FT	DOMAIN	446	494	CYTOPLASMIC (POTENTIAL).
FT	SITE	277	279	DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT	CARBOHYD	43	43	N-LINKED (GLCNAC...) (POTENTIAL).
FT	SEQUENCE	494 AA;	54194 MW;	A89204D3EA74BFBA CRC64;

Query Match 29.8%; Score 866; DB 1; Length 494;
 Best Local Similarity 38.9%; Pred. No. 2.1e-51;
 Matches 178; Conservative 90; Mismatches 188; Indels 2; Gaps 1;

QY	57	LVASLAGAPGSGFLYGNLSVNVNADPYIKAFYNESWERRGRDIDPDITLLMSVTVSI	116
DB	10	LFAISIAITIGSFQGYNTGVINAPEAIKDFLNYLTERSETPPSSVLLTSLMSLSVAI	69
QY	117	FAIGGLVGLIIVKMGKVLGRKHTLLANNNGFAISAAALMACSLQAGAFEMLIIVGRFIMGI	176
DB	70	FVGGMIGFSVGLFVNRFGRNWSLIIVNLAIAGGCLMGFCCKIAESVEMLIIGRLIIGL	129
QY	177	DGQVALSVLPMYLSRISPKRISGLSQVTAIFICIGVFTGQLGLPELLGKSTWPLYFG	236
DB	130	FCGLCTGFVPMYIGISPTALRGAGTNLQLGIVIGILVAQIFGLKVLGTEDLWPLLGG	189
QY	237	VIVPAVQVLLSLPFLDPSRYLLEKNEARAVKAFOTFLGKADVQOEVEEVLAEHVQ	296
DB	190	FTILPAIIQCAALPFCPSRFLINRKEEKAKIQLRWLTEDVAQDIQEMKDESIRM	249
QY	297	RSIRLSVLEILRAPVVRQVVTIVTMACYCLGLNAIWEFTNSISCKAGIPPAKIPYV	356
DB	250	SOEKQVTVLELFRAPNYRQPIIISIMLOISQOLSGINAVFYISTGIFKADGV--QEPVYA	307
QY	357	TLSTGGIETLAARVFGSLVIEHLGRPLLIIGFGLMGLFFGTITITITLQDHAPWVPYLSI	416
DB	308	TIGAGVNTIFTVSVFLVERAGRTLHLIGLGNVAFCSILMTISLLKONYSWMSFICI	367
QY	417	VGILAIASFCGPGGIPILTGEFFQSQRPAAFIAGTNNVLSNFAVGLLPFFIOKSL	476
DB	368	GAILVFAVFEIGPGPIPMFIVAELEFGQGPRAAMAVAGCSNWTSNFLVGLLPFSATFYL	427
QY	477	DTYCFVLPVATICITGAIXLYFVLPETKRTVAEISQAF	514
DB	428	GAYVIVFVTVFLVFWVFFFKVPETGRTEETRAP	465

Search completed: April 6, 2004, 10:17:27
 Job time : 20 secs

352 KIPYVTLSTGGIETLAASFGLVIEHLGRPL-LIGGFGMLGFFGTTLTITLQDHPW 410
324 ---YATIGAGVNVTVFTLVSVLLVERAGRTTLHLGLAGMCGCAI-LMTVALLLDERVA 379
411 VPYLSVIGLIALIASCSGPGGIPFILTGTGEFFQSORPAAFIAGTVNLSNPVGLLFP 470
380 MSYVSIVAIFGVFAFEIGPGIPWFIVAEVLSQGPFPAMAVAGFSNWTNFIIGMGFQ 439
471 FIOKSLDTYCLVLFATICTITGAIYLYFLVLPETKNRTYASISQAFSKRNKAYPPEEK 526
440 YVAEMGPYVLLFAVLLGLGFFFTFLRVPETGRFTDQISAAFHRTPSLLEQEVK 495

RESULT 15
GTR3_SHEEP
ID GTR3_SHEEP STANDARD; PRT; 494 AA.
AC P47843;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3, brain).
GN SLC2A3 OR GLUT3 OR GLUT-3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Coopeorth;
RX MEDLINE=96109471; PubMed=8653093;
RA Bennett B.L., Prosser C.G., Grigor M.R.;
RT "Isolation of cDNAs and tissue specific expression of ovine glucose
RT transporters.";
RL Biochem. Mol. Biol. Int. 37:9-16(1995).
RN [2]
RP SEQUENCE OF 88-248 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97392487; PubMed=9250701;
RA Currie M.J., Bassett N.S., Gluckman P.D.;
RT "Ovine glucose transporter-1 and -3: cDNA partial sequences and
RT developmental gene expression in the placenta.";
RL Placenta 18:393-401(1997).
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal
CC glucose transporter.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in brain.
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L39214; AAC41629.1; -;
CC EMBL; U89030; AAB49313.1; -;
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PRO0171; SUGTRNSPORT.
DR TIGRfam; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_2; 1.
DR Transmembrane; SUGAR transport; Transport; Glycoprotein;
KW Multigene family; Polymorphism; Disease mutation; Diabetes mellitus.
FT DOMAIN 1 24
FT TRANSMEM 25 45
FT DOMAIN 46 81
FT TRANSMEM 82 102
FT DOMAIN 103 111
FT TRANSMEM 112 132
FT DOMAIN 133 142
FT TRANSMEM 143 163
FT DOMAIN 164 171
FT TRANSMEM 172 192
FT DOMAIN 193 201
FT TRANSMEM 202 222
FT DOMAIN 223 287
FT TRANSMEM 288 308
FT DOMAIN 309 323
FT TRANSMEM 324 344
FT DOMAIN 345 353
FT TRANSMEM 354 374
FT DOMAIN 375 384
FT TRANSMEM 385 405
FT DOMAIN 406 417
FT TRANSMEM 418 438
FT DOMAIN 439 445
FT TRANSMEM 446 466
FT DOMAIN 467 509
FT SITE 295
FT
FT SITE 489 490
FT CARBOHYD 57 57
FT VARIANT 78 78
FT
FT VARIANT 383 383
FT VARIANT 385 385
FT
FT MUTAGEN 489 490
FT CONFLICT 151 154
FT SEQUENCE 509 AA; 54787 MW; 8E20CD97562C1BBF CRC64;
Query Match 30.1%; Score 874; DB 1; Length 509;
Best Local Similarity 38.7%; Pred. No. 6.2e-52;
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;
57 LVASLAGAFSGSFLYGYNLGVNNAAPTPIKAFYNESWERRHG---RPIDPDTLTLLWSV 112
24 LVLAFAVAVLGSQFGNIGVINAPQKVIQSYNETWLGQEGEGSSIPPGTLTLMAL 83
113 TVSIFAIQGGVGTLLIVKMGKVLGRKHTLLANNFSAIALMACSLQAGAFEMLIIVGRP 172
84 SVAIFSVGGMISFLIGIISQWLGRKEAMLVNVLAVLGSLMGLANAASVEMLIIGRF 143
173 IMGIDGVALSVTPMYLSEISPEKIRSLQGVTAIFICIVFGQGLLGLPELIGKSTWP 232
144 LIGAYSGLTSLGVPMYVGEIAPTHLRGALGTLLQLAIVIGILIAQVLGLESLLGTASLMP 203
233 YLFGVIVPAVQLLSLPFLPDSPRYLLLEKHEARAVKAFQTFGLKADYSQVVEVLAE 292
204 LLLGLTVLPALLQLVLPFCPESPRYLIITQNEGPARKSLKRLTGWADVSGVLAELKDE 263
293 SHVQRSLYSVLELLRAPYVRWQVTVIVTMACYQLCGNNAIWFTNYSIFGKAGI-PPA 351
264 KRKLRRERPLSLQLLGRSRRHQPLIIIAVVLQSLQSLGGINAVFYISTSIFFIAGVGQPA 323

InterPro; IPR003663; Sugar_transpt.
Pfam; PF00083; sugar tr; 1.
PRINTS; PRO0171; SUGTRNSPORT.
TIGRfam; TIGR00879; SP; 1.
PROSITE; PS50850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
Transmembrane; SUGAR transport; Transport; Glycoprotein;
KW Multigene family; Polymorphism; Disease mutation; Diabetes mellitus.
FT DOMAIN 1 24
FT TRANSMEM 25 45
FT DOMAIN 46 81
FT TRANSMEM 82 102
FT DOMAIN 103 111
FT TRANSMEM 112 132
FT DOMAIN 133 142
FT TRANSMEM 143 163
FT DOMAIN 164 171
FT TRANSMEM 172 192
FT DOMAIN 193 201
FT TRANSMEM 202 222
FT DOMAIN 223 287
FT TRANSMEM 288 308
FT DOMAIN 309 323
FT TRANSMEM 324 344
FT DOMAIN 345 353
FT TRANSMEM 354 374
FT DOMAIN 375 384
FT TRANSMEM 385 405
FT DOMAIN 406 417
FT TRANSMEM 418 438
FT DOMAIN 439 445
FT TRANSMEM 446 466
FT DOMAIN 467 509
FT SITE 295
FT
FT SITE 489 490
FT CARBOHYD 57 57
FT VARIANT 78 78
FT
FT VARIANT 383 383
FT VARIANT 385 385
FT
FT MUTAGEN 489 490
FT CONFLICT 151 154
FT SEQUENCE 509 AA; 54787 MW; 8E20CD97562C1BBF CRC64;
Query Match 30.1%; Score 874; DB 1; Length 509;
Best Local Similarity 38.7%; Pred. No. 6.2e-52;
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;
57 LVASLAGAFSGSFLYGYNLGVNNAAPTPIKAFYNESWERRHG---RPIDPDTLTLLWSV 112
24 LVLAFAVAVLGSQFGNIGVINAPQKVIQSYNETWLGQEGEGSSIPPGTLTLMAL 83
113 TVSIFAIQGGVGTLLIVKMGKVLGRKHTLLANNFSAIALMACSLQAGAFEMLIIVGRP 172
84 SVAIFSVGGMISFLIGIISQWLGRKEAMLVNVLAVLGSLMGLANAASVEMLIIGRF 143
173 IMGIDGVALSVTPMYLSEISPEKIRSLQGVTAIFICIVFGQGLLGLPELIGKSTWP 232
144 LIGAYSGLTSLGVPMYVGEIAPTHLRGALGTLLQLAIVIGILIAQVLGLESLLGTASLMP 203
233 YLFGVIVPAVQLLSLPFLPDSPRYLLLEKHEARAVKAFQTFGLKADYSQVVEVLAE 292
204 LLLGLTVLPALLQLVLPFCPESPRYLIITQNEGPARKSLKRLTGWADVSGVLAELKDE 263
293 SHVQRSLYSVLELLRAPYVRWQVTVIVTMACYQLCGNNAIWFTNYSIFGKAGI-PPA 351
264 KRKLRRERPLSLQLLGRSRRHQPLIIIAVVLQSLQSLGGINAVFYISTSIFFIAGVGQPA 323

Db 62 AFISAVLMGFKSGFEMILGRFIVGYCGUTTGFPVPMYGVESPTALRGALGTIHL 121
 QY 208 FIGVFTGQLLGPPELLGKSTWPLFGVIVPAVVQVLLSLPFLPDSRYLILLEKNEA 267
 Db 122 GIWVGLIAQVGLDSIMGNEELWPLLSVIFIPALLOCVLLPFCPSRFLINNEEN 181
 QY 268 RAVKAFOTFLGKADVSQVEVEVLAESHVQSRISLVSLLELLRAPYVRQVQVIVTMACY 327
 Db 182 RAKSVLKKLGTADVTRDLQEMKEESQRMREKKVITILEFRSAAYRQPIILIAVQLQSL 241
 QY 328 QLCGLNINWYTNISIFKAGIPPAKIPYVTLSTGGTETLAAVSGLVIEHLGRPL-LIG 386
 Db 242 QLSGINAVFYSTSIPEKAGV--QQPYATIGSGIVNTAFTVSLFVVERAGRTRHLIG 299
 QY 387 GFGLMGLFFGTLLTTLTQLQHPWPVYLSVGLILAIASCSGPGGIPFLITGFFQOQS 446
 Db 300 LAGVAGCAV-LMTIALALLSOLPMSYLSIVAFGVAFVEVGGPIFPWIVELSPSQP 358
 QY 447 RPAAFITAGTWLNSFAVGLLPFFIOKSLDLYCFLVFATICTITGAILYLYFLVPEKNRT 506
 Db 359 RPAIAVAGSNWTSNTVGMCFYVQELCGPYVFIITVLLVLFIFTYFKVPEIKGRT 418
 QY 507 YAEISQAF-----SKRNKAYPPEE 525
 Db 419 FDEIASGFRGGASQDKT--PEE 440
 RESULT 14
 GTR4 HUMAN STANDARD; PRT; 509 AA.
 ID GTR4 HUMAN STANDARD; PRT; 509 AA.
 AC P14672;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 4
 DE (Glucose transporter type 4, insulin-responsive).
 DE SLC2A4 OR GLUT4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=89255193; PubMed=2656669;
 RA Fukumoto H., Kayano T., Buse J.B., Edwards Y., Pilch P.F., Bell G.I.,
 RA Seino S.;
 RT "Cloning and characterization of the major insulin-responsive glucose
 RT transporter expressed in human skeletal muscle and other insulin-
 RN J. Biol. Chem. 264:7776-7779 (1989).
 RP SEQUENCE FROM N.A., AND VARIANT THR-385.
 RX MEDLINE=93012518; PubMed=1397719;
 RA Buse J.B., Yasuda K., Ley T.P., Seo T.S., Olson A.L., Pessin J.E.,
 RA Karam J.H., Seino S., Bell G.I.;
 RT "Human GLUT4/muscle-fat glucose-transporter gene. Characterization
 RT and genetic variation."
 RL Diabetes 41:1436-1445 (1992).
 RN [3]
 RP SEQUENCE OF 1-11 FROM N.A.
 RX MEDLINE=93366193; PubMed=7916714;
 RA Chiaramonte R., Martini R., Taramelli R., Comi P.;
 RT "Identification of the 5' end of the gene encoding a human insulin-
 RT responsive glucose transporter."
 RL Gene 130:307-308 (1993).
 RN [4]
 RP INTERACTION WITH DAXX, AND SUMOYLATION.
 RX MEDLINE=22028023; PubMed=11842083;
 RA Laloti V.S., Vergara Jauregui S., Pulido D., Sandoval I.V.;
 RT "The insulin-sensitive glucose transporter, GLUT4, interacts
 RT physically with Daxx. Two proteins with capacity to bind Ubc9 and
 RT conjugated to SUMO1."
 RL J. Biol. Chem. 277:19783-19791 (2002).

RN SUBCELLULAR LOCATION, AND MUTAGENESIS OF 489-LEU-LEU-490.
 RX MEDLINE=94131986; PubMed=8300557;
 RA Verhey K.J., Birnbaum M.J.;
 RT "A Leu-Leu sequence is essential for COOH-terminal targeting signal of
 RT GLUT4 glucose transporter in fibroblasts."
 RL J. Biol. Chem. 269:2353-2356 (1994).
 RN [6]
 RP VARIANT NIDDM ILE-383.
 RX MEDLINE=92012233; PubMed=1918382;
 RA Kusari J., Verma U.S., Buse J.B., Henry R.R., Olefsky J.M.;
 RT "Analysis of the gene sequences of the insulin receptor and the
 RT insulin-sensitive glucose transporter (GLUT-4) in patients with
 RT common-type non-insulin-dependent diabetes mellitus."
 RL J. Clin. Invest. 88:1323-1330 (1991).
 RN [7]
 RP VARIANT NIDDM ILE-383.
 RX MEDLINE=92097867; PubMed=1756912;
 RA Choi W.H., O'Rahilly S., Buse J.B., Rees A., Morgan R., Flier J.S.,
 RA Moller D.E.;
 RT "Molecular scanning of insulin-responsive glucose transporter (GLUT4)
 RT gene in NIDDM subjects."
 RL Diabetes 40:1712-1718 (1991).
 RN [8]
 RP VARIANT NIDDM ILE-383.
 RX MEDLINE=92394434; PubMed=1521731;
 RA O'Rahilly S., Krook A., Morgan R., Rees A., Flier J.S., Moller D.E.;
 RT "Insulin receptor and insulin-responsive glucose transporter (GLUT 4)
 RT mutations and polymorphisms in a Welsh type 2 (non-insulin-dependent)
 RT diabetic population."
 RL Diabetologia 35:486-489 (1992).
 CC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC -!- SUBUNIT: Binds to DAXX.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes
 CC primarily to the perinuclear region, undergoing continued
 CC recycling to the plasma membrane where it is rapidly
 CC reinternalized. The dileucine internalization motif is critical
 CC for intracellular sequestration.
 CC -!- TISSUE SPECIFICITY: Skeletal and cardiac muscles; brown and white
 CC fat.
 CC -!- PTM: Sumoylated.
 CC -!- DISEASE: Defects in SLC2A4 could be a cause of noninsulin-
 CC dependent diabetes mellitus (NIDDM) [MIM:125853]. Defects in
 CC SLC2A4 may be a cause of certain post-receptor defects in NIDDM.
 CC The variant in position Ile-383 is found in a small number of
 CC NIDDM patients, but seems not to be found in nondiabetic subjects.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M20747; AAA59189.1; -;
 CC EMBL; M91463; AAA52569.1; -;
 CC EMBL; X58489; CAA41399.1; -;
 CC PIR; A49158; A33801.
 CC Genew; HGNC:11009; SLC2A4.
 CC GK; P14672; -;
 CC MIM; 138190; -;
 CC MIM; 125853; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0005624; C:membrane fraction; TAS.
 CC GO; GO:0005355; F:glucose transporter activity; TAS.
 CC GO; GO:0005975; P:carbohydrate metabolism; TAS.
 CC GO; GO:0015758; P:glucose transport; TAS.
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR005828; Sub transporter.
 CC InterPro; IPR005829; Sug_transporter.

FT TRANSMEM 272 7 (POTENTIAL).
FT DOMAIN 293 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 307 8 (POTENTIAL).
FT DOMAIN 328 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 337 9 (POTENTIAL).
FT DOMAIN 338 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 338 10 (POTENTIAL).
FT DOMAIN 339 11 (POTENTIAL).
FT TRANSMEM 341 12 (POTENTIAL).
FT DOMAIN 342 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 343 11 (POTENTIAL).
FT DOMAIN 402 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 423 12 (POTENTIAL).
FT DOMAIN 430 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 451 492 DEFINES SUBSTRATE SPECIFICITY (BY
FT SITE 279 281 SIMILARITY).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 492 AA; 53962 MW; 05719AB061D0P67C CRC64;
Query Match 32.0%; Score 929.5; DB 1; Length 492;
Best Local Similarity 39.4%; Pred. No. 1e-55;
Matches 187; Conservative 98; Mismatches 179; Indels 11; Gaps 5;
QY 57 LVASLAGAGSFLGYNLNVNAPPYIKAFVNSWERRHGRPDPTLTLLWSVTYSI 116
Db 12 LMLVAGVGLSGQFGYNTGVINAPQKVEEFYNQWNRHRYGESIPSTLTLLWSLSVAI 71
QY 117 FAIGGLVGTLLVGMKVLGRKHTLLANNFSAISALLMACSLQAGAFEMLVGFRFMGI 176
Db 72 FSVGGMIGFSVGLFVNRFRGRNSMLMALLAFVSALVLMGFKSLGKSFEMLLIGRPIGV 131
QY 177 DGVVALSVLPMYLSSEIKRISGLQVTAIFICIGVFTGQLLGPPELLGKSTWPLYFG 236
Db 132 YCGLTTFGVPMYVGEVSPALRGALGTQLHQLGIWGLIAQVGLDSIMGNADLWPLLS 191
QY 237 VIVPVAWQLSLPFLPDSPRYLLEKNEARAVKAFQFLGKADVSQVEVEVLAESHVQ 296
Db 192 VIFIPALLQCLLPFCPSPRFLLINRENEARAKSVLKLGRGTADYTRDLOEMKEGRQM 251
QY 297 RSRLVSVLELLRAPYVRVQVTVIVTMACYOLCGLNALWFTNSTFGKAGIPPAKIPV 356
Db 252 MREKVTILELRSAPYROPILIAVVLQSLQSGINAVFYTSIFEKAGV--QPVVA 309
QY 357 TLSTGGIETLAAVFGSLVTEHLGRPL-LIGGFGMLGFFGTITLTLDHAPWPVYLS 415
Db 310 TTSGVLINTAFTVSLFVVERAGRRTLLHGLAGMACAV-LMTIALALLEQLPWSYLS 368
QY 416 IVGILAIASFCSGPGGPIITGTGFEFFQSQRPAAFIAGTWNWLNFAVGLLFFPIQKS 475
Db 369 IVAIFGVAFVFEVGPPIPWFIVAELFSQPRPAVAVAGFNSWTFVGMCFQVQEOL 428
QY 476 LDTYCLVLPATICITGAILVLYFVLPETKRTVABISQAF-----SKSNKAYPPEE 525
Db 429 CGPYVEIETVLLVLFETFEYKVPETKGTDEIASGFRQGSQSDXT--PEE 481
RESULT 13
GTRI_PIG STANDARD; PRT; 451 AA.
AC P20303;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 1
DE (Glucose transporter type 1, erythrocyte/brain) (Fragment).
GN SLC2A1 OR GLUT1
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=89302689; PubMed=2472815;
RA Weiler-Guettler H., Zinke H., Moeckel B., Frey A., Gassen H.G.;

FT "CDNA cloning and sequence analysis of the glucose transporter from
RL porcine blood-brain barrier.";
CC Biol. Chem. Hoppe-Seyler 370:467-473(1989).
CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be
CC responsible for constitutive or basal glucose uptake. Has a very
CC broad substrate specificity; can transport a wide range of aldoses
CC including both pentoses and hexoses.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes
CC primarily at the cell surface (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X17058; CAA34904.1; -;
DR PIR; S04223; S04223.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS50850; MPS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT NON_TER 1 1
FT DOMAIN <1 25 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 26 46 2 (POTENTIAL).
FT DOMAIN 47 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 75 3 (POTENTIAL).
FT DOMAIN 76 85 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 86 106 4 (POTENTIAL).
FT DOMAIN 107 114 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 115 135 5 (POTENTIAL).
FT DOMAIN 136 144 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 145 165 6 (POTENTIAL).
FT DOMAIN 166 230 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 231 251 7 (POTENTIAL).
FT DOMAIN 252 266 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 267 287 8 (POTENTIAL).
FT DOMAIN 288 296 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 297 317 9 (POTENTIAL).
FT DOMAIN 318 327 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 328 348 10 (POTENTIAL).
FT DOMAIN 349 360 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 361 381 11 (POTENTIAL).
FT DOMAIN 382 388 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 389 409 12 (POTENTIAL).
FT DOMAIN 410 451 CYTOPLASMIC (POTENTIAL).
FT SITE 238 240 DEFINES SUBSTRATE SPECIFICITY (BY
FT SIMILARITY).
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 451 AA; 49777 MW; E8F69C8A460B1028 CRC64;
Query Match 30.2%; Score 877.5; DB 1; Length 451;
Best Local Similarity 39.9%; Pred. No. 3.1e-52;
Matches 177; Conservative 92; Mismatches 164; Indels 11; Gaps 5;
QY 88 FYNESWERRHGRPDPTLTLLWSVTYSIFAIGLVGTLLVGMKVLGRKHTLLANNGF 147
Db 2 FYNQWTLHRYGESIPATLTLLWSLSVAIFSVGGMIGFSVGLFVNRFRGRNSMLMMLL 61
QY 148 AISALLMACSLQAGAFEMLVGFRFMGI 207

```
FT TRANSMEM 186 206 5 (POTENTIAL).
FT DOMAIN 207 271 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 307 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 308 328 8 (POTENTIAL).
FT DOMAIN 329 337 9 (POTENTIAL).
FT TRANSMEM 338 358 9 (POTENTIAL).
FT DOMAIN 359 371 10 (POTENTIAL).
FT TRANSMEM 372 392 10 (POTENTIAL).
FT DOMAIN 393 401 11 (POTENTIAL).
FT TRANSMEM 402 422 11 (POTENTIAL).
FT DOMAIN 423 429 12 (POTENTIAL).
FT TRANSMEM 430 450 12 (POTENTIAL).
FT DOMAIN 451 492 10 (POTENTIAL).
FT SITE 279 281 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CABOBYD 45 45 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 52 52 I -> Y (IN REF. 2).
FT CONFLICT 193 195 IFI -> VFV (IN REF. 1).
FT CONFLICT 357 360 LLER -> MQQ (IN REF. 4).
FT CONFLICT 403 403 A -> R (IN REF. 1).
SQ SEQUENCE 492 AA; 53934 MW; 4210646626F3735 CRC64;

Query Match 32.3%; Score 936.5; DB 1; Length 492;
Best Local Similarity 39.8%; Pred. No. 3.5e-56;
Matches 189; Conservative 96; Mismatches 179; Indels 11; Gaps 5;

QY 57 LVASLAGFGSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPTLTLLMSVTVSI 116
Db 12 LMLVGGAVLGSFGYNTGVNAPQKVFIEFYNTWNRHIGRIPSTLTLLMSLSVAI 71
QY 117 FAIGLVGTLLVQMGKVLGRKHTLLANGFAISAALLMACSLQAGAFEMLVGRFMGI 176
Db 72 FSVGGMIGSFVGLFVNFGRNLSMNMNLLAFVAAVLMSFKLGKSFEMILGRFTIGV 131
QY 177 DGVNLSVLMVLSISPKETGSLGQVTAIFICIGVTGQLLGLPELLGKESWYPLFG 236
Db 132 YCGLTGTGVPMVGVSVSTALRGALGTLHQLGIVVGLIAQVGLDSMGNADLWPLLS 191
QY 237 VIVPAVQQLSLPLPDSRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQ 296
Db 192 VIFIPALQCLLLPSPSPRPLLNNEHRAKSVLKLRTGADVTDLDQEMKEEQM 251
QY 297 RSIRLVSVLELLRAPYRWQVTVITVACVQLCGLNAIFVYVTSIFGKAGIPPAKIPYV 356
Db 252 MREKVTILELFRSAYQPIIAVLQSLQSGINAVFYVSTSEKAGV--QQPYA 309
QY 357 TLSTGIGETLAVFSGLVIEHLGRPL-LIGFGMLGFFGTLITLTDHAPWVPLS 415
Db 310 TIGSGIVNTAFTVWSLFFVVERAGRTLHLGLAGMAGCAV-LMTTALLERLPMSYLS 368
QY 416 IVGILAITASFCGPGGIPFILTGEFFQSQRPRAFIITAGTVNLSNFAVGLLPFIQKS 475
Db 369 IVAIFGVAFPEVFGPIFWFVAELFSQGPRAIAVAGFSNWTSNFVGMCFQYVEQL 429
QY 476 LTYCFLVPATCITGAIYLVVLVLPETKRTVAETISQAF-----SKRNKAYPPEE 525
Db 429 CGPYVFIITVLLVLFPIFTYFKVETKRTFDEIASGRQCGASQSDKT--PEE 481

RESULT 12
GRL1_RAT
ID GRL1_RAT STANDARD; PRT; 492 AA.
AC P11167;
DT 01-JUL-1989 (Rel. 11, Created)
DT 21-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 1
DE (Glucose transporter type 1, erythrocyte/brain).
GN SLC2A1 OR GLUT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

FT	CONFLICT	152	152	/FTID=VAR 013286.
FT	SEQUENCE	492 AA	54117 MW	F -> L (IN REF. 3).
SQ	SEQUENCE	492 AA	54117 MW	3B3472479834429D CRC64;
Query Match				
Best Local Similarity 39.6%; Pred. No. 3 5e-56;				
Matches 188; Conservative 98; Mismatches 178; Indels 11; Gaps 5;				
QY	57	LVASLAGAFSGSFLGYNSVNVNAPPTYPYKAFYNESWERRHORIPDPTLTLLWSVTVSI	116	
DB	12	LMLAVGVALGSLQFGYNTGVINAPQKVIEEFYNTQWVHRYGESILPTLTLLWSLVAI	71	
QY	117	FAIGLIVGLTVKMGKVLGRKHTLLANNFPAISALLMACSLQAGAPEMLIVGFINGI	176	
DB	72	FSVGGMIGSVGLFVNRFRGRNSMLMMNLLAFVSAVLMGFKLQKSPKSPMLILGRFIIGV	131	
QY	177	DGVALSVLPKYLSEISPKIEGSLQGVTAIFICIGVFTGQLLPELLEKSTWPLYFG	236	
DB	132	YCGLTGTFYVGVESPTAFKALGTLHQLGVLIVGILLAQVFGDLSINGKDLNPLLS	191	
QY	237	VTVPAVQVQLSLPFDSPRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAESHVQ	296	
DB	192	IFIPALLQCVLPFCFSPRFLINRNEENRAKSVLKLRTADVTHDLQEMKEESRQM	251	
QY	297	RSIRLSVLELRAPYRVQWVTVVTVMACYQGLNATWFTNSIFGKAGIPPAKIPV	356	
DB	252	MREKVTILELSPSPAYROPILIAVVLQSLQSGINAVFYSTISFERAGV--QOPVYA	309	
QY	357	TUSTGTGIELAAVFGSLVIEHGRRL-LIGFGLMGLFFGTLITLTLDHAPWPVYLS	415	
DB	310	TIGSGIVNTAFTVSLFVVERAGRTLHLGLAGMACAI-LMTIALALLEQLPMMVYLS	368	
QY	416	IVGILAIASFSGCGGIPILTGEFFQSQRPAAFIAGTVNLSNFAVGLLFFPIQKS	475	
DB	369	IVAIFGVFAFFVGPPIFWFVVAELFSGQPRPAATAVAGFNSWTFVGMCFQVVEQL	428	
QY	476	LDTYCFVLVATICITGAILYFVLPETKNTYAEISQAF-----SKRNKAYPPEE	525	
DB	429	CGPVFIITVILVLFITFYFKVETKGTPTDEIASGRQGSQSDKT--PEE	481	

RESULT 11

ID	GTR1_MOUSE	STANDARD;	PRT;	492 AA.
AC	P17809; Q61608;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Solute carrier family 2, facilitated glucose transporter, member 1			
DE	(Glucose transporter type 1, erythrocyte/brain) (Glt1).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
[1]	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RC	MEDLINE=89240694; PubMed=2654938;			
RA	Kaesner K.H., Christy R.J., McLenth J.C., Braiterman L.T.,			
RA	Cornelius P., Pekala P.H., Lane M.D.;			
RT	"Sequence, tissue distribution, and differential expression of mRNA			
RT	for a putative insulin-responsive glucose transporter in mouse 3T3-L1			
RT	adipocytes."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:3150-3154 (1989).			
[2]	SEQUENCE FROM N.A.			
RC	MEDLINE=90274408; PubMed=2190533;			
RA	Reed B.C., Shade D., Alperovich F., Vang M.;			
RT	"3T3-L1 adipocyte glucose transporter (HepG2 clasp): sequence and			
RT	regulation of protein and mRNA expression by insulin,			
RT	differentiation, and glucose starvation."			
RL	Arch. Biochem. Biophys. 279:261-274 (1990).			

[3]	SEQUENCE OF 151-237 FROM N.A., AND DEVELOPMENTAL STAGE.			
RA	STRAIN=C57BL/6; TISSUE=Embryo;			
RC	MEDLINE=93170163; PubMed=1289053;			
RA	Smith D.E., Gridley T.;			
RT	"Differential screening of a PCR-generated mouse embryo cDNA library:			
RT	glucose transporters are differentially expressed in early			
RT	postimplantation mouse embryos."			
RL	Development 116:555-561 (1992).			
[4]	SEQUENCE OF 357-463 FROM N.A., AND DEVELOPMENTAL STAGE.			
RA	STRAIN=CD-1; TISSUE=Embryo;			
RC	MEDLINE=92111400; PubMed=1765007;			
RA	Hogan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,			
RA	Jenkins N.A., Thorens B., Schultz G.A.;			
RT	"Glucose transporter gene expression in early mouse embryos."			
RL	Development 113:363-372 (1991).			
CC	-!- FUNCTION: Facilitative glucose transporter. This isoform may be			
CC	responsible for constitutive or basal glucose uptake. Has a very			
CC	broad substrate specificity; can transport a wide range of aldoses			
CC	including both pentoses and hexoses.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes			
CC	primarily at the cell surface (by similarity).			
CC	-!- DEVELOPMENTAL SPAGE: Levels decline 3-fold between days 7.5 and			
CC	12.5 of gestation. At 7.5 dpc, expressed more strongly in			
CC	extraembryonic tissues than in the embryo proper. Expressed in			
CC	amion, chorion, and ectoplacental cone. In the yolk sac,			
CC	expressed more strongly in the mesoderm layer than the ectoderm.			
CC	Expression fairly widespread in the embryo at 8.5 dpc, but by 10.5			
CC	dpc, expression is down-regulated and observed in the eye and the			
CC	spinal cord.			
CC	-!- SIMILARITY: Belongs to the sugar transporter family. Glucose			
CC	transporter subfamily.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M23384; AAA37752.1; -			
DR	EMBL; M22998; AAA37707.1; -			
DR	EMBL; X69697; CAA49367.1; -			
DR	EMBL; S77924; AAB20846.2; -			
DR	PIR; S09705; S09705.			
DR	MGI; MGI:95755; SLC2a1.			
DR	GO; GO:0005622; C:intracellular; IDA.			
DR	GO; GO:0005886; C:plasma membrane; IDA.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR005828; Sub_transporter.			
DR	InterPro; IPR005829; Sug_transporter.			
DR	InterPro; IPR003663; Sugar_transpt.			
DR	Pfam; PF00083; sugar tr; 1.			
DR	PRINTS; PR00171; SUGTRANSPT.			
DR	TIGRfam; TIGR00879; SP; 1.			
DR	PROSITE; PS00850; MFS; 1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
KW	Transmembrane; Sugar transport; Transp; Glycoprotein;			
KW	Multigene family.			
FT	DOMAIN	1	12	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	13	33	1 (POTENTIAL).
FT	DOMAIN	34	66	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	67	87	2 (POTENTIAL).
FT	DOMAIN	88	95	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	96	116	3 (POTENTIAL).
FT	DOMAIN	117	126	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	127	147	4 (POTENTIAL).
FT	DOMAIN	148	155	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	156	176	5 (POTENTIAL).
FT	DOMAIN	177	185	EXTRACELLULAR (POTENTIAL).

Yu W., Gibbs R.A.;
Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.
[4]
SEQUENCE OF 394-423 FROM N.A.
RP
RC TISSUE=Articular cartilage;
RA Neama G., Richardson S., Sell S., Carter S., Mobasheri A.;
RT "Molecular characterization and cloning of glucose transporters in
human articular chondrocytes.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
[5]
VARIANT GLUT1 DEFICIENCY ILE-310.
RX MEDLINE=99242249; PubMed=10227690;
RA Klepper J., Wang D., Fischbach J., Vera J.C., Jarjour I.T.,
O'Driscoll K.R., De Vivo D.C.;
RA "Defective glucose transport across brain tissue barriers: a newly
recognized neurological syndrome.";
RL Neurochem. Res. 24:587-594(1999).
[6]
VARIANTS GLUT1 DEFICIENCY PHE-66; LEU-126; LYS-146; GLU-256 AND
RP TRP-333.
RX MEDLINE=20438345; PubMed=10980529;
RA Wang D., Kranz-Eble P., De Vivo D.C.;
RA "Mutational analysis of GLUT1 (SLC2A1) in Glut-1 deficiency
syndrome.";
RL Hum. Mutat. 16:224-231(2000).
[7]
ERRATUM.
RP Wang D., Kranz-Eble P., De Vivo D.C.;
RL Hum. Mutat. 16:527-527(2000).
[8]
VARIANT GLUT1 DEFICIENCY HIS-126.
RX MEDLINE=21487108; PubMed=11603379;
RA Brockmann K., Wang D., Korenke C.G., von Moers A., Ho Y.-Y.,
Pasqual J.M., Kuang K., Yang H., Ma L., Kranz-Eble P., Fischbach J.,
Hansfeld F., De Vivo D.C.;
RA "Autosomal dominant Glut-1 deficiency syndrome and familial
epilepsy.";
RL Ann. Neurol. 50:476-485(2001).
[9]
VARIANT GLUT1 DEFICIENCY ASP-91.
RX MEDLINE=21065632; PubMed=11136715;
RA Klepper J., Willemssen M., Verrips A., Guerteen E., Herrmann R.,
Kutzick C., Floercken A., Voit T.;
RA "Autosomal dominant transmission of GLUT1 deficiency.";
RL Hum. Mol. Genet. 10:63-68(2001).
CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be
responsible for constitutive or basal glucose uptake. Has a very
broad substrate specificity, can transport a wide range of aldoses
including both pentoses and hexoses.
CC CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes
primarily at the cell surface [By similarity].
CC CC
CC -!- TISSUE SPECIFICITY: Expressed at variable levels in many human
tissues.
CC CC
CC -!- DISEASES: Defects in SLC2A1 are the cause of autosomal dominant
GLUT1 deficiency syndrome [MIM:608777]. This disease causes a
defect in glucose transport across the blood-brain barrier. It is
characterized by infantile seizures, delayed development, and
acquired microcephaly.
CC CC
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).

EMBL; K03195; AAA52571.1; --
DR EMBL; M20653; AAB61084.1; --
DR EMBL; AF070544; AAC28635.1; --
DR EMBL; AY034633; AAAS6795.1; --

[illegible]

GTR1 HUMAN

011160

DT-100

DT 10-000

התאריך: 11.11.2011

DE 37

SO
HOTO

OC Mamma

11

५३००५८	५५
--------	----

RA Mueck

RT "Sequ."

101 102

RP SEQUENCES

RA Fukum.

gr trans.

TABLE 1

RD SEQUES

Db 191 FIFVPALQOICILPFAPESEFRFLINRNEENKAKSVLKKLGGTTDVSSDLQEMKESESRQM 250
QY 297 RSRLSVLELLRAPVVRQWVTVITMACYQCLGLNAIWFYTNIGFGKAGIPPAKIPYV 356
Db 251 MREKKTIVLELFRSPMYRQPIIAIVLQLSQSGINAVFYSTISFEKSGV--EQPVYA 308
QY 357 TLSTGGIETLAAVFGSLVIEHLGRRPL-LIGFGELMGLFFGTLITLTLQDHAPWVPLS 415
Db 309 TIGSGVNTAFTVSVLFVVERAGRRTHLIGLAGMAGCAI-LMTIALTLDDQPMWSYLS 367
QY 416 IVCILAIISFSCGPGGIFILTGFEFFQSQSORPAAFIAGTVNLSNFAVGLLEPFIQKS 475
Db 368 IVAIFGVAFPEIGPGPIPMFIVAEFLPSQGPRAFAVAGLSNWTSNFIVGMPQIAOL 427
QY 476 LDYCYFLVATICITGAIVLYFVLPETKRTYAEISQAF-----SKRNKAYPEE 525
Db 428 CGSYVFIIFVLVLFFIIFYFKVPETKGTDFEIAIYRQGGASQDKT--PDE 480

RESULT 8

GTR3_CHICK
ID GTR3_CHICK STANDARD; PRT; 496 AA.
AC P28568;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Soluce carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3) (CER-GT3).
GN SLC2A3 OR GLUT3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91342646; PubMed=1875932;
RA White M.K.; Rall T.B.; Weber M.J.;
RT "Differential regulation of glucose transporter isoforms by the arc
RT oncogene in chicken embryo fibroblasts.";
RL Mol. Cell. Biol. 11:4448-4454(1991).
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal
CC glucose transporter.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M37785; AAA48662.1; -
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGTRNSPORT.
CC TIGRFAMs; TIGR00879; SP; 1.
CC PROSITE; PS50850; MFS; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
CC Transmembrane; Sugar_transport; Transpt; Glycoprotein;
KW Multigene family.
KW DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 27 1 (POTENTIAL).
FT DOMAIN 28 65 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 66 86 2 (POTENTIAL).
FT DOMAIN 87 94 CYTOPLASMIC (POTENTIAL).
FT

FT TRANSMEM 95 115 3 (POTENTIAL).
FT DOMAIN 116 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 146 4 (POTENTIAL).
FT DOMAIN 147 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 6 (POTENTIAL).
FT DOMAIN 206 270 7 (POTENTIAL).
FT TRANSMEM 271 291 8 (POTENTIAL).
FT DOMAIN 292 306 9 (POTENTIAL).
FT TRANSMEM 307 327 10 (POTENTIAL).
FT DOMAIN 328 335 11 (POTENTIAL).
FT TRANSMEM 336 356 12 (POTENTIAL).
FT DOMAIN 357 367 10 (POTENTIAL).
FT TRANSMEM 368 388 11 (POTENTIAL).
FT DOMAIN 389 398 12 (POTENTIAL).
FT TRANSMEM 399 419 10 (POTENTIAL).
FT DOMAIN 420 426 11 (POTENTIAL).
FT TRANSMEM 427 447 12 (POTENTIAL).
FT DOMAIN 448 496 12 (POTENTIAL).
FT SITE 278 280
FT CARBOHYD 44 44
FT SEQUENCE 496 AA; 54174 MW; 75B3C0F61A7A92A5 CRC64;
SQ
Query Match 32.5%; Score 942.5; DB 1; Length 496;
Best Local Similarity 39.3%; Pred. No. 1.4e-56;
Matches 193; Conservative 94; Mismatches 187; Indels 17; Gaps 5;
QY 49 RDKWSCSLVSLAGAFSGSSFLYGNLSVNVNAPPYTKAFVNESWERRHGRIPDPDLTL 108
Db 4 KKKITASLIYAVSAAIG-SLQFGYNTGVINAPEKIIQAFYNTLSORSGETISPELLTS 62
QY 109 LWSVTVSIFAAGLVGTLIVRMIGVLRKHTLLANNNGFAISAAALLMACSIQAQAFEMLI 168
Db 63 LWSUVAIFSGVMIGSVSVSLFFNRFGRNSMLLVNVLAFAGGALMALSIAKAVEMLI 122
QY 169 VGRFMIDGGVALSVLPMYLSSEISPKIEIRGLSGQVTAIFTCIGVFTQQLGLPELLGKE 228
Db 123 IGRFIIGFCGLCTGCFVPMYISEVPTSLRGAFGLNQLGIWGLVVAIIFGLEGIMGTE 182
QY 229 STWPLYFGVIVPAVQVLLSPFLPDSPPYLLLEKHNEARAVKAFQIFLGKADYSQVEE 288
Db 183 ALWPELLGFTIVPAVLQCVALLFCPESPRFLINKKEEKAQTVLQKLRGTQDVSDISE 242
QY 289 VLAESHVQSRISLNSVLSLLRAPYRVQVTVITMACYQCLGLNAIWFYNTSIFGKAGI 348
Db 243 MKEESAKMSQEKATVLELFRSPNTRQPIIISITLQLSQSGINAVFYSTIGFPERAGI 302
QY 349 PPAKIPYVTLSTGGIETLAAVFGSLVIEHLGRRPILLIGFGMLGLFFOTLITLLODHA 408
Db 303 --TQPVYATIGAGVNTVTVSVLELVERAGRTLHIVLGGMAVCAAVMTIALAKE-- 358
QY 409 PWVPYLSIVGILAIISAFSCGPGGIFILTGFEFFQSQSORPAAFIAGTVNLSNFAVGLL 468
Db 359 KWRIVISIVATGTFVALFEIGPGPIPMFIVAEFLPSQGPRAFAVAGCSNWTSNFLVGM 418
QY 469 PFFIOKSLDTCFLVPAICITGAIVLYFVLPETKRTYAEISQAFSKRNKAYPEEKID 528
Db 419 FFYAEKLCGPYVFLIFLVLFFIIFYFKVPETKGTDFEISRGF-----EQQVE 469
QY 529 SAVTDAPASSP 539
Db 470 ---TSPPSSPP 477
RESULT 9
GTR1_RABIT STANDARD; PRT; 492 AA.
ID GTR1_RABIT
AC P13355;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT

primarily at the cell surface (By similarity).
-!- SIMILARITY: Belongs to the sugar transporter family. Glucose transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L07300; AAB02037.1; --
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub transporter.
InterPro; IPR005829; Sug transporter.
InterPro; IPR003663; Sug transept.
Pfam; PF00083; sugar tr_1.
PRINTS; PR00171; SUGRTRNSPORT.
TIGRFAMs; TIGR00879; SP; 1.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
KW DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 12 32 1 (POTENTIAL).
FT DOMAIN 33 65 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 66 86 2 (POTENTIAL).
FT DOMAIN 87 94 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 95 115 3 (POTENTIAL).
FT DOMAIN 116 125 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 126 146 4 (POTENTIAL).
FT DOMAIN 147 154 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 184 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 185 205 6 (POTENTIAL).
FT DOMAIN 206 270 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 271 291 7 (POTENTIAL).
FT DOMAIN 292 306 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 307 327 8 (POTENTIAL).
FT DOMAIN 328 336 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 337 357 9 (POTENTIAL).
FT DOMAIN 358 369 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 370 390 10 (POTENTIAL).
FT DOMAIN 391 400 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 401 421 11 (POTENTIAL).
FT DOMAIN 422 428 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 429 449 12 (POTENTIAL).
FT DOMAIN 450 490 CYTOPLASMIC (POTENTIAL).
FT SITE 177 179 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 44 44 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 490 AA; 54086 MW; 63BB176812D06104 CRC64;
Query Match 32.6%; Score 946.5; DB 1; Length 490;
Best Local Similarity 40.6%; Pred. No. 7.4e-57;
Matches 193; Conservative 95; Mismatches 176; Indels 11; Gaps 5;
Qy 57 LVASLAGAFSGSFLYGVNLSVVNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVSI 116
Db 11 LMLAVGAVLSGLQFGVNTGVINRPOKVEDFNHTWLYREDFIPATLTLWSLSVAI 70
Qy 117 FAIGGLVGTLLIVKMGKVLGRKHTLLANNGFAISALLMACSLQAGAFEMLIYGRFMGI 176
Db 71 FSVGGMIGSFVGLFVYNRFGRRNSMLNLAFLAFLVLMGFSKVALSFEMILGRFIIGL 130
Qy 177 DGGVALSVLPMYLSPEIKRISGLQVTAIFTCIGVFTQQLGLPELLGKSTWYLPFG 236
Db 131 YSGLTTGFVPMYGVESPTALRGALGTTHQGLVILGILIAQVFGSLDLMGNSLWPLLLG 190
Qy 237 VIVPAVAVQLLSPLPDSRYLLLEKHNARAVKAFQTFGLKADYSQEEVEVLAEHVQ 296

DOMAIN 423 429 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 430 450 12 (POTENTIAL).
FT DOMAIN 451 492 CYTOPLASMIC (POTENTIAL).
FT SITE 279 281 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 492 AA; 54131 MW; D83605D0F13D7252 CRC64;
Query Match 32.7%; Score 949.5; DB 1; Length 492;
Best Local Similarity 40.0%; Pred. No. 4.6e-57;
Matches 190; Conservative 100; Mismatches 174; Indels 11; Gaps 5;
Qy 57 LVASLAGAFSGSFLYGVNLSVVNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVSI 116
Db 12 LMLAVGAVLSGLQFGVNTGVINRPOKVEDFNHTWLYREDFIPATLTLWSLSVAI 71
Qy 117 FAIGGLVGTLLIVKMGKVLGRKHTLLANNGFAISALLMACSLQAGAFEMLIYGRFMGI 176
Db 72 FSVGGMIGSFVGLFVYNRFGRRNSMLNLAFLAFLVLMGFSKVALSFEMILGRFIIGV 131
Qy 177 DGGVALSVLPMYLSPEIKRISGLQVTAIFTCIGVFTQQLGLPELLGKSTWYLPFG 236
Db 132 YSGLTTGFVPMYGVESPTALRGALGTTHQGLVILGILIAQVFGSLDLMGNSLWPLLLS 191
Qy 237 VIVPAVAVQLLSPLPDSRYLLLEKHNARAVKAFQTFGLKADYSQEEVEVLAEHVQ 296
Db 192 VIFIPALQOCHLLPPCPSPFLNNEHRKSVLKURGTADVTRDLQEMKESRQM 251
Qy 297 RSIRLVSVLELLRAPYVQVVTIVTWACYQLCGLNAIWFTYNSIFGKAGIPPAKIPYV 356
Db 252 MREKVTILELFRSAAYQPILIAVLQLSQGLSINAVFYVSISPEKAGV--QQPYVA 309
Qy 357 TLSTGGIETLAAPFGLVIEHLGRPL-LIGFGMLGLFFGTLLITLQDHAPWVPLS 415
Db 310 TIGSGIVNTAFTVSVLFFVVERAGRTHLIGLAGMAGCAV-LMTIALALLERLPWMSYLS 368
Qy 416 IVGLIATIASCSGGGIPFILTGGFFQSQRPAAFIAGTVNLSFAVGLLPFFIOKS 475
Db 369 IVAIFGFVAFVEVGPGIPFWFVIAELFSQGRPAIAVAGFSNMTSNFVGMCFQYVEQL 428
Qy 476 LDYTCFLVFATICITGAYLVPLPETKNRTVAEISQAF----SKRNKAYPPBE 525
Db 429 CGPVYFIITLVLLVLFPIFTFKVETKGRFDEIASGFRGGASQSDKT--PEE 481
RESULT 7
GTRI_CHICK STANDARD; PRT; 490 AA.
AC P46896;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 1
DE (Glucose transporter type 1) (GTL1).
GN SLC2A1 OR GLUT1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96157892; PubMed=8589457;
RA Wagstaff P., Kang H.Y., Mylott D., Robbins P.J., White M.K.;
RT "Characterization of the avian GLUT1 glucose transporter:
RT differential regulation of GLUT1 and GLUT3 in chicken embryo fibroblasts";
RL Mol. Biol. Cell 6:1575-1589(1995).
CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be responsible for constitutive or basal glucose uptake. Has a very broad substrate specificity; can transport a wide range of alcohols including both pentoses and hexoses (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes

DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfam; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
KW DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 32 1 (POTENTIAL).
FT DOMAIN 33 67 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 68 88 2 (POTENTIAL).
FT DOMAIN 89 100 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 101 121 3 (POTENTIAL).
FT DOMAIN 122 131 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 132 152 4 (POTENTIAL).
FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 5 (POTENTIAL).
FT DOMAIN 182 192 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 193 213 6 (POTENTIAL).
FT DOMAIN 214 276 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 277 297 7 (POTENTIAL).
FT DOMAIN 298 314 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 315 335 8 (POTENTIAL).
FT DOMAIN 336 341 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 342 362 9 (POTENTIAL).
FT DOMAIN 363 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 10 (POTENTIAL).
FT DOMAIN 389 411 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 412 432 11 (POTENTIAL).
FT DOMAIN 433 438 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 439 459 12 (POTENTIAL).
FT DOMAIN 460 501 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 50 50 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 501 AA; 55524 MW; 3696501DE11DC5C1 CRC64;

Query Match 37.9%; Score 1099.5; DB 1; Length 501;
Best Local Similarity 45.1%; Pred. No. 3.4e-67; Indels 3; Gaps 1;
Matches 214; Conservative 94; Mismatches 163;

QY 56 LLVASIAGAFGSFLYGVNLSVNVNAPTPIKAFYNESWERRRGRIPDPTLLLSVTVS 115
DB 16 LALATLIIAAGFSFGYGVNVAVNSPSEFMQFYNDYDREENIESFTLLLSLVTS 75

QY 116 IFAIGLVGTIVVMIGKVLGRKHTLLANNFSAALMACSLOQAFEMLIIVGRFPM 175
DB 76 MFPPGGFISGLVGVNVLNKLGRKALLFNFIPLPAIFMGCSQIAQSPELLIISRLV 135

QY 176 IDGVALSIVLPMYLSISPKETRGSLGQVTAIPICIGVTGQLGLPELLGKSTWPIYLF 235
DB 136 ICAGISSNVVPMYLGELAPKRLRGLGVVPQLPFITVGLVLAQLFGLRLSLANEDGNPVL 195

QY 236 GVIVVPAVQLLSLPFLPDSRVLLEKHEBAVKAFOFELGKADVSQVEEVLAEHV 295
DB 196 GLTGVPAGLQLLPPFPSPRYLLIOKKDEAAERALTQIRGWKDVHLEMEIRKDEBA 255

QY 296 QRSIRLVSVLELRLAPYVQVWVTVVTVMACVQLCGINAIWFYNTSIFGKAGIPPAKIPY 355
DB 256 EXAAGFISVWKLFTMQLSRQLISMIVLMAGQQLSGVNIYYADQIYLSAGVKSDVQY 315

QY 356 VTLSGGIETLAIVFSLGVIHGRPLIGFGMLGFPFTLTITLTQDHAPVPIYLS 415
DB 316 VTAGTGANVVMFTLITFVVELWGRFRLLLVGFTSTCLIACLVLTALALQNTISWNPYIS 375

QY 416 IVGILAIIASFCGPGGPIILTGEPFQSQSORPAAFIAGTVNLSNFAVGLLFPQKS 475
DB 376 IVCVIVVIVGHALGPSIPALLITEIFLOSSRPAAYMIGSVHLSNFTVGVIFPFIQMG 435

QY 476 LDFYCLVLPATICITGAIYLVFLPETKNTYAEISQAFSRKKNK---APPPEK 526
DB 436 LGPYSFIIFATICLFTIYIYFVVPETKGRFTIELIQLFNMKNKVSQDVTPKKEE 489

RESULT 5
GTL1_HUMAN STANDARD; PRT; 496 AA.
AC Q9BYW1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 11
DE (Glucose transporter type 11) (Glucose transporter type 10).
GN SLC2A11 OR GLUT11 OR GLUT10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Heart;
RA Doegge H., Bocianski A., Scheepers A., Axer H., Eckel J., Joost H.-G.;
RT "Two novel members of the family of sugar transport facilitators:
RT genomic organization and predominant expression in lymphocytes (GLUT9)
RT and muscle (GLUT10).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21616996; PubMed=11741323;
RA Sasaki T., Minoshima S., Shiohama A., Shintani A., Shimizu A.,
RA Asakawa S., Kawasaki K., Shimizu N.;
RT "Molecular cloning of a member of the facilitative glucose
RT transporter gene family, GLUT11 (SLC2A11) and identification of
RT transcription variants.";
RL Biochem. Biophys. Res. Commun. 289:1218-1224(2001).
CC -!- FUNCTION: Facilitative glucose transporter (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=GLUT11;
CC IsoId=Q9BYW1-1; Sequence=Displayed;
CC Name=2; Synonyms=GLUT11-s;
CC IsoId=Q9BYW1-2; Sequence=VSP_006293, VSP_006294, VSP_006295;
CC -!- TISSUE SPECIFICITY: Expressed in heart and skeletal muscle.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTERS SUBFAMILY.
CC -!- CAUTION: Has been described as GLUT10 in literature, but this
CC gene name has already been used for SLC2A10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ271290; CAC29020.1; -;
CC EMBL; AB049214; BAB68410.1; -;
CC Genew; HGNC:14239; SLC2A11.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGTRNSPORT.
CC TIGRfam; TIGR00879; SP; 1.
CC PROSITE; PS00850; MFS; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transp. sugar transport; Transmembrane; Glycoprotein;
KW Multigene family; Alternative splicing.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 32 1 (POTENTIAL).
FT DOMAIN 33 67 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 68 88 2 (POTENTIAL).
FT DOMAIN 89 100 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 101 121 3 (POTENTIAL).
FT DOMAIN 122 131 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 132 152 4 (POTENTIAL).
FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 5 (POTENTIAL).
FT DOMAIN 182 192 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 193 213 6 (POTENTIAL).
FT DOMAIN 214 276 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 277 297 7 (POTENTIAL).
FT DOMAIN 298 314 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 315 335 8 (POTENTIAL).
FT DOMAIN 336 341 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 342 362 9 (POTENTIAL).
FT DOMAIN 363 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 10 (POTENTIAL).
FT DOMAIN 389 411 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 412 432 11 (POTENTIAL).
FT DOMAIN 433 438 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 439 459 12 (POTENTIAL).
FT DOMAIN 460 501 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 50 50 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 501 AA; 55524 MW; 3696501DE11DC5C1 CRC64;

RA Kasahara T., Kasahara M.,
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochalasin B-sensitive carrier. Seems to function
 CC primarily as a fructose transporter.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in small intestine, and at much
 CC lower levels in kidney, skeletal muscle, and adipose tissue.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L05195; AAA02627.1; -;
 CC EMBL; D13871; BAA02983.1; -;
 CC EMBL; D28562; BAA05912.1; -;
 CC PIR; I53268; I53268.
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR005828; Sub_transporter.
 CC InterPro; IPR005829; Sug_transporter.
 CC InterPro; IPR003663; Sugar_transpt.
 CC Pfam; PF00083; sugar_tr; 1.
 CC PRINTS; PR00171; SUGETRNSPORT.
 CC TIGRams; TIGR00879; SP; 1.
 CC PROSITE; PS00850; MFS; 1.
 CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 CC Transmembrane; Sugar transport; Transp; Glycoprotein;
 CC Multigene family.
 FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 12 32 1 (POTENTIAL).
 FT DOMAIN 33 71 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 72 92 2 (POTENTIAL).
 FT DOMAIN 93 99 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 100 120 3 (POTENTIAL).
 FT DOMAIN 121 131 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 132 152 4 (POTENTIAL).
 FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 161 181 5 (POTENTIAL).
 FT DOMAIN 182 191 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 192 212 6 (POTENTIAL).
 FT DOMAIN 213 276 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 277 297 7 (POTENTIAL).
 FT DOMAIN 298 314 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 315 335 8 (POTENTIAL).
 FT DOMAIN 336 342 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 343 363 9 (POTENTIAL).
 FT DOMAIN 364 367 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 368 388 10 (POTENTIAL).
 FT DOMAIN 389 411 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 412 432 11 (POTENTIAL).
 FT DOMAIN 433 438 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 439 459 12 (POTENTIAL).
 FT DOMAIN 460 502 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 164 164 V -> A (IN REF. 1).
 FT CONFLICT 285 285 A -> T (IN REF. 2).
 FT CONFLICT 452 452 T -> S (IN REF. 2).
 FT CONFLICT 479 479 K -> N (IN REF. 1).
 SQ SEQUENCE 502 AA; 55543 MW; 871FEDEFALE7684D CRC64;

Query Match 38.3%; Score 1112.5; DB 1; Length 502;
 Best Local Similarity 44.9%; Pred. No. 4.5e-68;
 Matches 218; Conservative 96; Mismatches 167; Indels 5; Gaps 2;

QY 56 LVLASLAGAGSGFLYGNLSVNVNATPVYIKAFYNESWERRHGRIPDPTLLWSVTVS 115
 DB 16 LALATFLAAGFSFGYGNVAVNSPESFMQOQFYNDTYDRNKENIESFTLTLWSLTVS 75

QY 116 IFAGGLVGTLLVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYVGRFMG 175
 DB 76 MFPGFGFGLSLVGVFLVNNLGRKGLALLFNNI-FSLIPAILMGCSKIAXSFEIILASRLVG 135
 QY 176 IDGVALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPGLLCKESTWYLF 235
 DB 136 ICAGISSNVVPMYLGELAPKNLRGALGVVPQLFITVGLVLAQLFGLRSVLASEGWPILL 195
 QY 236 GVIVVPAVQVLLSLRFLPDSRYLLEKHNARAVKAFOTELGKADVSQEEVEVLASHV 295
 DB 196 GUTGPDAGLQQLLLFFPESRYLLIQKNESAELALQTLRGKQVDMEMEERKEDEA 255
 QY 296 QRSIRLVSVLELLRAPPYRWQVTVVIMACYQLGLNAINWFTYNSIFGKAGIPAKIPY 355
 DB 256 EKAAAGFISVWKLFRMQSLRWQLISTVLMAGQQLSGVNAIYVYADQIYLSAGVKSNDVQY 315
 QY 356 VTLSGGTETLAASVSGVIEHLGRPLLIIGGGLMGLFGFTLTITITLQDHPAPVYLS 415
 DB 316 VTAGTGANVFMVTMTVFVVELWGRNLLI-GFSTCLTACIVLVALALQNTISWMIYVS 375
 QY 416 IVGLILAIASFQSGFGGIPFILTGEFFQSQORPAFIAGTVNMLSNFAVGLLFPFIQKS 475
 DB 376 IVCVIVYVIGHAVGSPIPALFITEIFLOSSEPRAYMIGSVHMLSNFVGLIFFPQVG 435
 QY 476 LPTYCFLVPATICITGAILYLVLPETKNRYVAISQAFSGKNC---AYPEEKIDSAVT 532
 DB 436 LQPSYFIIPAIICLLTIVFMVPTKGRTFVBEINQIFAKKNKVSVDYVP--EKEEKELN 493
 QY 533 DAPASS 538
 DB 494 DLPAPT 499
 RESULT 4
 GTR5_MOUSE
 ID GTR5_MOUSE STANDARD; PRT; 501 AA.
 AC Q9WV38;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 5
 GN (Glucose transporter type 5, small intestine) (Fructose transporter).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=20207171; PubMed=12031501;
 RA Kwon O., Levine M., Burant C.F.;
 RT "Cloning and functional characterization of the mouse fructose
 RT transporter, GLUT5.";
 RL Biochim. Biophys. Acta 1576:191-197(2002).
 CC -!- FUNCTION: Cytochalasin B-sensitive carrier. Seems to function
 CC primarily as a fructose transporter (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTERS SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF161071; AAD42235.1; -;
 CC MGD; MGI:1928369; Slc2a5.
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR005828; Sub_transporter.

lower levels in kidney, skeletal muscle, and adipose tissue.
-!- MASS SPECTROMETRY: MW=54973.98; METHOD=MALDI.
-!- SIMILARITY: Belongs to the sugar transporter family. Glucose transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch>).

EMBL; M55531; AAB52570.1; -
EMBL; U11843; AAB60641.1; -
EMBL; U05344; AAB60641.1; JOINED.
EMBL; U11839; AAB60641.1; JOINED.
EMBL; U11840; AAB60641.1; JOINED.
EMBL; U11841; AAB60641.1; JOINED.
EMBL; U11842; AAB60641.1; JOINED.
EMBL; BC001692; AAB01692.1; -
EMBL; BC001820; AAB01820.1; -
PIR; A36629; A36629.
PIR; G02864; G02864.
Genew; HGNC:11010; SLC2A5.
MIM; 138230; -
GO; GO:0005886; C:plasma membrane; TAS.
GO; GO:0005353; F:fructose transporter activity; TAS.
GO; GO:0005355; F:glucose transporter activity; TAS.
GO; GO:0005975; P:carbohydrate metabolism; TAS.
GO; GO:0015755; P:fructose transport; TAS.
GO; GO:0015758; P:glucose transport; TAS.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub transporter.
InterPro; IPR005829; Sug transporter.
InterPro; IPR003683; Sugar_transpt.
Pfam; PF00083; sugar_tr; 1.
PRINTS; PR00171; SUGTRNSPORT.
TIGRfams; TIGR00879; SP; 1.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
Transmembrane; Sugar transport; Transport; Glycoprotein;
Multigene family.
DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 33 1 (POTENTIAL).
FT DOMAIN 34 68 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 69 88 2 (POTENTIAL).
FT DOMAIN 90 98 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 99 119 3 (POTENTIAL).
FT DOMAIN 120 120 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 121 141 4 (POTENTIAL).
FT DOMAIN 142 161 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 162 182 5 (POTENTIAL).
FT DOMAIN 183 192 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 193 213 6 (POTENTIAL).
FT DOMAIN 214 277 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 278 298 7 (POTENTIAL).
FT DOMAIN 299 319 8 (POTENTIAL).
FT TRANSMEM 319 339 8 (POTENTIAL).
FT DOMAIN 340 342 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 343 363 9 (POTENTIAL).
FT DOMAIN 364 368 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 369 389 10 (POTENTIAL).
FT DOMAIN 390 412 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 413 433 11 (POTENTIAL).
FT DOMAIN 434 439 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 440 460 12 (POTENTIAL).
FT DOMAIN 461 501 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 12 44 RLTLVLAALIAAFSSFOYGNVAVNSPAL ->
P -> G (IN REF. 2).
CONFLICT 202 202

SQ SEQUENCE 501 AA; 54974 MW; 090812P529C7B857 CRC64;
Query Match 38.7%; Score 1122.5; DB 1; Length 501;
Best Local Similarity 46.4%; Pred. No. 9.4e-69;
Matches 222; Conservative 92; Mismatches 155; Indels 9; Gaps 3;
Qy 56 LLVASIAGAGSSFLGYNLSVNVATPYIKAFVNESWERRHGRPTDPTLTLLWSVTVS 115
Db 17 LALATLIAAFSSFOYGNVAVNSPALLMQQYNYTYGRTGEFDEDFPLTLWSVTVS 76
Qy 116 IFAIGGLVGLTVKMGKVLGRKHTLLANNFSAISALLMACSLQAGAPEMLIVGRFIMG 175
Db 77 MFPFGFGIGLLVGPLVKNFGRKALLFNFIIVPAILMGCSRVATSFELIISRLVG 136
Qy 176 IDGGVALSVLPMYLSSEISPKIRSGISQVTAIFICIGVFTGQLLGLPELLGKESTWPLYF 235
Db 137 ICAGVSSNVVPMVLGELAPXNLGALGVVPLQTLITVILVAQIFGLRNLILANVDGWPILL 196
Qy 236 GVIIVPAVWOLLSLPLPDSPRVLLLEKNEARAVAFQTFELGKADVSOQVEEVLAEHV 295
Db 197 GLTGVPAAQLQLLPFPFSPRYLLIQKDEAAAKKALQTLRGWDSVDREVAIRQDEA 256
Qy 296 QRSIRLVNLELLRAPVVRVQVTVITMACYOLCGLNATVFTNSTFGKAGIPPAKIPY 355
Db 257 EKAAGFISVLKLFRRSLRWQLLSIIVLMGGQQSGVNAIYYADQIYLSAGVPEEHVQY 316
Qy 356 VTLSTGGIE---TLAAVFSGLVTEHLGRRRPLLIIGGFLMGLFFGTITITLLODHAPWVP 412
Db 317 VTAGTAVNVVMTFCVFP---VVELLGRLLLLGFSICLIACCVLTAALALQTVSWMP 373
Qy 413 YLSIVGILAITASPCSGPGGIPFLTGEFPQQSORPAAFIAGTVNLSHPAVGLLPFFI 472
Db 374 YISIVCVISYVIGHALGSPIPALLITEIFLQSSRPSAFVGVGSHVLSNFTVLGIFPFI 433
Qy 473 QKSLDTYCFVPAITCITGAIYFLVLPETQRTYARISQAFSKRNK---AYPPEEKI 527
Db 434 QEGLGYSFIVFAVICLLTIYFLIYFETKAKTFIEINQIFTKQNKVSEYVPEKEEL 491
RESULT 3
GTRS RAT
ID GTRS RAT STANDARD; PR7; 502 AA.
AC P43427;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 5
DE (Glucose transporter type 5, small intestine) (Fructose transporter).
GN SLC2A5 OR GLUT5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Jejunum;
RX MEDLINE=93325725; PubMed=8333543;
RA Rand E.B., Depauli A.M., Davidson N.O., Bell G.I., Burant C.F.;
RT "Sequence, tissue distribution, and functional characterization of
RT the rat fructose transporter GLUT5."
RL Am. J. Physiol. 264:G1169-G1176(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Jejunum;
RX MEDLINE=94008761; PubMed=8404647;
RA Inukai K., Asano T., Katagiri H., Ishihara H., Anai M.,
RA Fukushima Y., Tsukuda K., Kikuchi M., Yazaki Y., Oka Y.;
RT "Cloning and increased expression with fructose feeding of rat jejunal
RT GLUT5."
RL Endocrinology 133:2009-2014(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Intestine;

DOMAIN	1	51	CYTOPLASMIC (POTENTIAL).
TRANSMEM	52	72	1 (POTENTIAL).
DOMAIN	73	107	EXTRACELLULAR (POTENTIAL).
TRANSMEM	108	128	2 (POTENTIAL).
DOMAIN	129	140	CYTOPLASMIC (POTENTIAL).
TRANSMEM	141	161	3 (POTENTIAL).
DOMAIN	162	171	EXTRACELLULAR (POTENTIAL).
TRANSMEM	172	192	4 (POTENTIAL).
DOMAIN	193	200	CYTOPLASMIC (POTENTIAL).
TRANSMEM	201	221	5 (POTENTIAL).
DOMAIN	222	231	EXTRACELLULAR (POTENTIAL).
TRANSMEM	232	252	6 (POTENTIAL).
DOMAIN	253	316	CYTOPLASMIC (POTENTIAL).
TRANSMEM	317	337	7 (POTENTIAL).
DOMAIN	338	354	EXTRACELLULAR (POTENTIAL).
TRANSMEM	355	375	8 (POTENTIAL).
DOMAIN	376	381	CYTOPLASMIC (POTENTIAL).
TRANSMEM	382	402	9 (POTENTIAL).
DOMAIN	403	415	EXTRACELLULAR (POTENTIAL).
TRANSMEM	416	436	10 (POTENTIAL).
DOMAIN	437	451	CYTOPLASMIC (POTENTIAL).
TRANSMEM	452	472	11 (POTENTIAL).
DOMAIN	473	478	EXTRACELLULAR (POTENTIAL).
TRANSMEM	479	499	12 (POTENTIAL).
DOMAIN	500	540	CYTOPLASMIC (POTENTIAL).
CARBOHYD	90	90	N-LINKED (GLCNAC. . .) (POTENTIAL).
VARIANT	25	25	R -> G.
FT			/FTID=VAR_012157.
FT	282	282	V -> I.
FT			/FTID=VAR_012158.
FT	350	350	L -> P.
FT			/FTID=VAR_012159.
SEQUENCE	540 AA; 58816 MW; E593661A0DDC0FE8	CRC64;	

Query Match 93.7%; Score 2719; DB 1; Length 540;
Best Local Similarity 99.4%; Pred. No. 1e-176;
Matches 530; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 MAPKONENSEKELGIWPLTDDTSHAGPGPGRALLCDHLRSYVPGGRRRRKDWCSLLVAS 60

Query Match 93.7%; Score 2719; DB 1; Length 540;
Best Local Similarity 99.4%; Pred. No. le-176;
Matches 530; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1	MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLCEDHLRSGVPGRRRRKDWSCSLLVAS	60	QY
1	MARKQNRNSKELGLVPLTDDTSHARPPGGRALLCEDHLRSGVPGRRRRKDWSCSLLVAS	60	Db
61	LACAFGSSFLYGYNLSVVNAPTPIKAFYNESHERHGRPIDPDTITLLMSVTVSIFAIG	120	QY
61	LACAFGSSFLYGYNLSVVNAPTPIKAFYNESHERRGRPIDPDTITLLMSVTVSIFAIG	120	Db
121	GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAILLMACSLQAGAFEMLIIVGRFIMGIDGGV	180	QY
121	GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAILLMACSLQAGAFEMLIIVGRFIMGIDGGV	180	Db
181	ALSVLPMYLSSEISPKETRGSIGQVTAIFICIGVFTGQLLPELLGKESWPLYLFGVIVV	240	QY
181	ALSVLPMYLSSEISPKETRGSIGQVTAIFICIGVFTGQLLPELLGKESWPLYLFGVIVV	240	Db
241	PAVVQLLSLFPDPSPRYLLEKNEARAVKAFOTFLGKADVSOEVEVLAEHVQRSIR	300	QY
241	PAVVQLLSLFPDPSPRYLLEKNEARAVKAFOTFLGKADVSOEVEVLAEHVQRSIR	300	Db
301	LVSVLELLRAPYVRQWQVTVVMWACYQLCGLNAINWFYTNISIFGKAGIPPAKIPVYTLST	360	QY
301	LVSVLELLRAPYVRQWQVTVVMWACYQLCGLNAINWFYTNISIFGKAGIPPAKIPVYTLST	360	Db
361	GGIETLAAVPSGLVIEHLGRPELLIGGFLMGLPFGTITITLTLDQHAPWPYLSIVGIL	420	QY
361	GGIETLAAVPSGLVIEHLGRPELLIGGFLMGLPFGTITITLTLDQHAPWPYLSIVGIL	420	Db
421	AIIASFCSGPGGIPFLITGEFFQOSORPAAFIIAGTVNWLNSNFAVGLLFPFIQKSLDNYC	480	QY
421	AIIASFCSGPGGIPFLITGEFFQOSORPAAFIIAGTVNWLNSNFAVGLLFPFIQKSLDNYC	480	Db
481	FLVPATCITCGAIVLYFVLPTKQRTYAEISQAFSKRNKAYPPEBKIDSAVTD	533	QY
481	FLVPATCITCGAIVLYFVLPTKQRTYAEISQAFSKRNKAYPPEBKIDSAVTD	533	Db

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:00:52 ; Search time 18 Seconds
(without alignments)
1628.639 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903
Sequence: 1 MARKQNRNRSKELGLVPLTDD.....NTAMTQAAATTTATKKEHPL 563

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2719	93.7	540	1 GTR9 HUMAN	Q9NRMO homo sapien
2	1122.5	38.7	501	1 GTR5 HUMAN	P22732 homo sapien
3	1112.5	38.3	502	1 GTR5 RAT	P43427 rattus norv
4	1099.5	37.9	501	1 GTR5 MOUSE	Q9WV38 mus musculus
5	999	34.4	496	1 GTR11 HUMAN	Q9BW41 homo sapien
6	949.5	32.7	492	1 GTR1 BOVIN	P27674 bos taurus
7	946.5	32.6	490	1 GTR1 CHICK	P46896 gallus gall
8	942.5	32.5	496	1 GTR3 CHICK	P28568 gallus gall
9	940.5	32.4	492	1 GTR1 RABIT	P13355 oryctolagus
10	936.5	32.3	492	1 GTR1 HUMAN	P11166 homo sapien
11	929.5	32.0	492	1 GTR1 MOUSE	P17809 mus musculus
12	877.5	30.2	451	1 GTR1 RAT	P11167 rattus norv
13	877.5	30.2	451	1 GTR1 PIG	P20303 sus scrofa
14	874	30.1	509	1 GTR4 HUMAN	P14672 homo sapien
15	866	29.8	494	1 GTR3 SHEEP	P47843 ovis aries
16	864	29.8	494	1 GTR3 BOVIN	P58352 bos taurus
17	854	29.4	509	1 GTR4 RAT	P19357 rattus norv
18	853	29.4	509	1 GTR4 MOUSE	P14142 mus musculus
19	844.5	29.1	493	1 GTR3 MOUSE	P32037 mus musculus
20	841	29.0	495	1 GTR3 CANFA	P47842 canis fami
21	841	29.0	509	1 GTR4 BOVIN	Q27994 bos taurus
22	830	28.6	496	1 GTR3 HUMAN	P11169 homo sapien
23	823.5	28.4	486	1 GTR5 RABIT	P46408 oryctolagus
24	816.5	28.1	493	1 GTR3 RAT	Q07647 rattus norv
25	772	26.6	522	1 GTR2 RAT	P12336 rattus norv
26	767.5	26.4	523	1 GTR2 MOUSE	P14246 mus musculus
27	750.5	25.9	533	1 GTR2 CHICK	Q90592 gallus gall
28	737.5	25.4	390	1 GTR1 SHEEP	P79365 ovis aries
29	695.5	24.0	524	1 GTR2 HUMAN	P11168 homo sapien
30	679.5	23.4	400	1 GTR3 RABIT	Q9XSC2 oryctolagus
31	480	16.5	488	1 YB91 YEAST	P38142 saccharomyc
32	457	15.7	204	1 GTR1 BOVIN	P58353 bos taurus
33	438	15.1	461	1 CSBC_BACSU	P46333 bacillus su

34 413 14.2 464 1 ARAE_BACSU
35 411.5 14.2 629 1 MYCT_HUMAN
36 410.5 14.1 507 1 GTR6 HUMAN
37 409.5 14.1 464 1 GALP_ECOLI
38 407.5 14.0 547 1 GTR1 LEIDO
39 400 13.8 468 1 GLCP_SYNY3
40 398 13.7 763 1 RGT2 YEAST
41 396.5 13.7 477 1 GTR8 MOUSE
42 394.5 13.6 477 1 GTR8 HUMAN
43 389.5 13.4 510 1 HEXE_RICCO
44 389 13.4 540 1 HXTD_YEAST
45 389 13.4 592 1 HXT5_YEAST

ALIGNMENTS

RESULT 1

GTR9_HUMAN
ID_GTR9_HUMAN STANDARD; PRT; 540 AA.
AC Q9NRMO;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 9
DE (Glucose transporter type 9).
GN SLC2A9 OR GLUT9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS GLY-25; ILE-282 AND PRO-350.
RC TISSUE=Kidney;
RX MEDLINE=20318624; PubMed=10860667;
RA Phay J.E., Hussain H.B., Moley J.F.;
RT "Cloning and expression analysis of a novel member of the facilitative glucose transporter family, SLC2A9 (GLUT9).";
RL Genomics 66:217-220(2000).
CC -!- FUNCTION: Facilitative glucose transporter (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney followed by liver; also detected in placenta, lung, blood leukocytes, heart and skeletal muscle.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE TRANSPORTERS SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AF210317; AAF85942.1; -
CC Genew; HGNC:13446; SLC2A9.
CC MIM; 606142; -
CC GO; GO:0016021; C:integral to membrane; NAS.
CC GO; GO:0005351; F:sugar porter activity; NAS.
CC GO; GO:0015758; P:glucose transport; NAS.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transpt.
CC Pfam; PF00083; sugar tr; 1.
CC PRINTS; PR00171; SUGTRNSPORT.
CC TIGRfam; TIGR00879; SP; 1.
CC PROSITE; PS00850; MFS; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Glycoprotein; Multigene family; Polymorphism.

Db 438 QYVADRMGPVYVLLFAVLLGFFFTFLKVPETGRKTFDQISAAFRTPSLLEQEVK 494

Job time : 22 secs

RESULT 15

S53322
fructose transport protein GLUT5, intestinal - rabbit
N;Alternate names: facilitative glucose transporter GLUT5; fructose transporter GLUT5
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 20-Jun-2000
C;Accession: S53322; S71039
R;Miyamoto, K.; Tatsumi, S.; Morimoto, A.; Minami, H.; Yamamoto, H.; Sone, K.; Taketani, Biochem. J. 303, 877-883, 1994
A;Title: Characterization of the rabbit intestinal fructose transporter (GLUT5).
A;Reference number: S53322; MUID:95071304; PMID:7980458
A;Accession: S53322
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-487 <MIY>
A;Cross-references: GB:D26482
R;Miyamoto, K.I.
submitted to the EMBL Data Library, January 1994
A;Reference number: S71039
A;Accession: S71039
A;Molecule type: mRNA
A;Residues: 1-424,426-487 <MIW>
A;Cross-references: EMBL:D26482; NID:G624911; PIDN:BAA05492.1; PID:G633998
C;Genetics:
A;Gene: GLUT5
C;Superfamily: glucose transport protein
C;Keywords: glycoprotein; sugar transport; transmembrane protein

Query Match 28.9%; Score 838; DB 2; Length 487;
Best Local Similarity 39.0%; Pred. No. 7.2e-59;
Matches 193; Conservative 91; Mismatches 181; Indels 30; Gaps 13;

QY 45 GRRRKDWCSLLVA--SLAGAFSGSFLYGYNLVVNAPTPIYKAFYNESWERRHGRPID 102
DB 5 GQEKKEGRITLVALRTLIAAFGSSFOYAVNVVNCSPSELMTEFYNDYVDTGELID 64
QY 103 PDTLTLLWSVTGIFPAIGLGVTLIVKMGKVLGRKHTLLANNNGPAISAILMACSIOAG 162
DB 65 EPELTLLWSVTVMFSPSGGFGAGLLVGPLVNFGRKGLLFFNNFISIVPAIIMGCKRVAR 124
QY 163 APFMLIVGPIGMDIGGVALSIVLPMYLSISPKETRGSLGQVTAIFIGVFTGQLGLP 222
DB 125 SFELIILRLVIGICAGVSNVVPYVGLAPNIRGALGVESQLFTILGILVAQIFGL- 183
QY 223 ELGKESWVPLFGVIVVPAVQVLSLPFLPDSRYLLEKNEARAYKAFQTLGKADV 282
DB 184 RSIRQKQWPIILLGLTGPA--ACPPFPSPRYLLIGQPRCRQ-KALQSLRGWDSV 240
QY 283 SQVEEVLAEHVORIRLVSVLELLRAPYVWQVTVIVTMACYQLGLNAINWFTNSI 342
DB 241 DRELEIRREDEAARAAGLVSVRALCAMGLAWQLISVVPLM-WOQLSGVNAIYY-DQI 298
QY 343 FGKAGIPP--AKIPYVTLTGTTETLAAVPSGLVIEHLRRPLLIGFGMLPFGTLTI 400
DB 299 Y-----LSPDLDTQYVTAATGAVNVLMVCTVVFVESWARLILL-LGFSPLAPTCCLTA 353
QY 401 TLTLDQHPWVYLSVIGILAIASFCSPGPGGIFILTGEFFQSQRPAAFIAGTVNWL 460
DB 354 ALALQDVTSMWPYISVICIVVIGHAIGP-AIRSLYT-EIFLQSGRPPTW--NGQVHVL 409
QY 461 SNFVGLLFPFIQKSLDTCFLVFATICTIGAIYLVLPETKNTYAEISQAFSKNKA 520
DB 410 SNFTVGLVFPFLQWALGLYSFIIFGVACLSLTTVYFLIVPETKGSFIEIIRFIRNKV 469
QY 521 -----YYPE 524
DB 470 EVSPDREELKDFPPD 484

Search completed: April 6, 2004, 10:18:58

Query Match 29.4%; Score 854; DB 2; Length 509;
Best Local Similarity 37.9%; Pred. No. 4.1e-60;
Matches 181; Conservative 99; Mismatches 185; Indels 12; Gaps 6;
QY 57 LVASLAGAFSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDP-----TLTLWS 111
DB 24 LVAVFSAVLGSLQFGYNGVINAQKVIEQSYNATWLGROG-PGGPDSIPQGTTLTWA 82
QY 112 VTVSFAIGLVGTIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGR 171
DB 83 LSVAFSVGGMISFLIGIISQWLGRKRAMLANNVLAIVGLGALMLANAAASYEILILGR 142
QY 172 FIMGIDGGVALSVLPMYLSLSPKIEIRSGISLGOVTAIFIGVFTGQLLGLPELIGKSTW 231
DB 143 FLIAGYSGTSLGVPVYGEIAPTHLRGALGTNLQLAIVIGILVAQVIGLESMLGTATLW 202
QY 232 PYLFGVIVPAVQVLLSPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVL 291
DB 203 PLLALTVPALLQQLLPPCPSPRYLYIIRNLEGPARKSLKRLTGADVSDALAEKLD 262
QY 292 ESHVQSRIRLVSVLELLRAPYRVQVTVVITMACYQVLCGLNAINFYNSIFGKAGI-PP 350
DB 263 EKKLERERPLSLQLLGSRTTHROPLIIAVVLQSLQSGINAVFYSTISFELAGVEQP 322
QY 351 AKIPYVTLSTGGIETLAUVFSGLVIEHLGRRL-LIGGFLMGLFFGTTLTTLTLODHAP 409
DB 323 A---YATIGAGVNVTVTLVSVLLVERAGRRTLHLGLAGMCGCAI-LMTVALLLRRVP 378
QY 410 WPYLSIVGILAIASFCSPGGIPFLITGEPFQOSQORPAAFIAGTVNMLSNFAVGLIF 469
DB 379 SMSYVSIIVAFGFAFEIGPGPIWFIIVAEFSGQPRPAANAVAGSNWTCNFVGNMF 438
QY 470 PTIOKSLDYCYFLVFATIGTCITGAIYLYFVLPETKRTYAEISQAFSKRNKAYPPEEK 526
DB 439 QYVADANGPVYFLFAVLLGFFITFLVPEVTRGRTFQISATFRTSLLQEYK 495
RESULT 13
A41751
Glucose transport protein 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 24-Sep-1999
C:Accession: A41751; S18128
R:Nagamatsu, S.; Kornhauser, J.M.; Burant, C.F.; Seino, S.; Mayo, K.E.; Bell, G.I.
J. Biol. Chem. 267, 467-472, 1992
A:Title: Glucose transporter expression in brain. cDNA sequence of mouse GLUT3, the brain
A:Reference number: A41751; MUID:92112695; PMID:1730609
A:Accession: A41751
A:Molecule type: mRNA
A:Residues: 1-493 <NAG>
A:Cross-references: GB:X61093; NID:g51088; PIDN:CAA43406.1; PID:g51089
C:Genetics:
A:Gene: GLUT3
C:Superfamily: glucose transport protein
C:Keywords: glycoprotein; transmembrane protein
F/43/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 29.1%; Score 844.5; DB 2; Length 493;
Best Local Similarity 38.8%; Pred. No. 2.2e-59;
Matches 179; Conservative 95; Mismatches 182; Indels 5; Gaps 4;
QY 55 SLIVASLAGAFSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDPDTLLMSVTV 114
DB 9 SLVFAVTVATIG-SFQFGYNGTVINAPETILKDFLNYTLERLEDLPSEGLLTALWSLCV 67
QY 115 SIFAIGLVTLIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGRFTM 174
DB 68 AIFSVMGIMGSFVGVFNFRGRNMLNLAIIAGCLMGFAKIAESVEMILGRLLII 127
QY 175 GIDGGVALSVLPMYLSLSPKIEIRSGISLGOVTAIFIGVFTGQLLGLPELIGKSTW 234
DB 128 GIFCGICTGVPMIGEVSTAIRGAFGTNLQNGIVGVILVAQIFGLDFILSSEELWPG 187

QY 235 FGIVVPAVVQLLSLSPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVLAE 294
DB 188 LGLTILPAIQSAALPPCPSPRYLLEKHNARAVKAPQTFGLKADVSQVEEVLAE 247
QY 295 VQSRIRLVSVLELLRAP-YRVQVTVVITMACYQVLCGLNAINFYNSIFGKAGI 353
DB 248 RMSQEKQVTVLELFRSNVYQVLLISIVLQLS-QQLSGINAVFYSTIGIFKADG 304
QY 354 PYVTLSTGGIETLAUVFSGLVIEHLGRRL-LIGGFLMGLFFGTTLTTLTLODHAP 413
DB 305 IYATIGAGVNVTVTLVSVLLVERAGRRTLHLGLGMAVCSVFMISLLKXDYEAM 364
QY 414 LSVIGILAIASFCSPGGIPFLITGEPFQOSQORPAAFIAGTVNMLSNFAVGLIF 473
DB 365 VCIIVAILIYVAFEIGPGPIWFIIVAEFSGQPRPAANAVAGSNWTCNFVGNMF 424
QY 474 KSLDYCYFLVFATIGTCITGAIYLYFVLPETKRTYAEISQAF 514
DB 425 AVLGAIVFIIFAAPLIFLFTFPPKVPETKRTFEDIARAF 465
RESULT 14
B30310
Glucose transport protein GT2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Dec-1989 #sequence_revision 08-Dec-1989 #text_change 04-Sep-1998
C:Accession: B30310
R:Kaestner, K.H.; Christy, R.J.; McLenithan, J.C.; Braiteman, L.T.; Cornelius, P.; Pek
Proc. Natl. Acad. Sci. U.S.A. 86, 3150-3154, 1989
A:Title: Sequence, tissue distribution, and differential expression of mRNA for a putat
A:Reference number: A30310; MUID:89240694; PMID:2654938
A:Accession: B30310
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-508 <KAE>
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein
Query Match 28.9%; Score 839.5; DB 2; Length 508;
Best Local Similarity 37.7%; Pred. No. 5.8e-59;
Matches 180; Conservative 99; Mismatches 185; Indels 13; Gaps 7;
QY 57 LVASLAGAFSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDP-----TLTLWS 111
DB 24 LVAVFSAVLGSLQFGYNGVINAQKVIEQSYNATWLGROG-PGGPDSIPQGTTLTWA 82
QY 112 VTVSFAIGLVGTIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGR 171
DB 83 LSVAFSVGGMISFLIGIISQWLGRKRAMLANNVLAIVGLGALMLANAAASYEILILGR 142
QY 172 FIMGIDGGVALSVLPMYLSLSPKIEIRSGISLGOVTAIFIGVFTGQLLGLPELIGKSTW 231
DB 143 FLIAGYSGTSLGVPVYGEIAPTHLRGALGTNLQLAIVIGILVAQVIGLESMLGTATLW 202
QY 232 PYLFGVIVPAVQVLLSPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVL 291
DB 203 PLLALTVPALLQQLLPPCPSPRYLYIIRNLEGPARKSLKRLTGADVSDALAEKLD 262
QY 292 ESHVQSRIRLVSVLELLRAPYRVQVTVVITMACYQVLCGLNAINFYNSIFGKAGI-PP 350
DB 263 EKKLERERPLSLQLLGSRTTHROPLIIAVVLQSLQSGINAVFYSTISFELAGVEQP 322
QY 351 AKIPYVTLSTGGIETLAUVFSGLVIEHLGRRL-LIGGFLMGLFFGTTLTTLTLODHAP 409
DB 323 A---YATIGAGVNVTVTLVSVLLVERAGRRTLHLGLAGMCGCAI-LMTVALLLRRVP 378
QY 410 WPYLSIVGILAIASFCSPGGIPFLITGEPFQOSQORPAAFIAGTVNMLSNFAVGLIF 469
DB 379 SMSYVSIIVAFGFAFEIGPGPIWFIIVAEFSGQPRPAANAVAGSNWTCNFVGNMF 437
QY 470 PTIOKSLDYCYFLVFATIGTCITGAIYLYFVLPETKRTYAEISQAFSKRNKAYPPEEK 526

QY 387 GFLMGLFTGTTITITLQDHAPWVYLSIVGILAIISFSCGPGGIPFILTGTFFQSQ 446
DB 300 LAGMACCAV-LMTIALALLEQLPMMYSIVAIFGVFAFFVFGPIFWFAELFSQGP 358
QY 447 RPAAFIITAGVWLNFAVGLLFPFIQKSLDNYCFLVPATICITGAILYXFLVLPETKRT 506
DB 359 RPAAIAGVWNTSNFVGMCFQVBEQCGPYFIITVLLVFFIYFKVPETKRT 418
QY 507 YAEISQAF-----SKENKAYPPEE 525
DB 419 FDEIASGFQGGASQSDKT--PEE 440
RESULT 11
A33801
muscle-fat glucose-transporter GLUT4 - human
N;Alternate names: glucose transport protein, insulin-regulated, solute carrier family 2
C;Species: Homo sapiens (man)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 24-Sep-1999
C;Accession: A49158; A33801; I55358; I37425
R;Buse, J.B.; Yasuda, K.; Lay, T.P.; Seo, T.S.; Olson, A.L.; Pessin, J.E.; Karam, J.H.;
Diabetes 41, 1436-1445, 1992
A;Title: Human GLUT4/muscle-fat glucose-transporter gene. Characterization and genetic v
A;Reference number: A49158; MUID:93012518; PMID:1397719
A;Accession: A49158
A;Molecule type: DNA
A;Residues: 1-509 <BUS>
A;Experimental source: fetal liver
A;Note: sequence extracted from NCBI backbone (NCBIN:117256, NCBIIP:117257)
R;Fukunoto, H.; Kayano, T.; Buse, J.B.; Edwards, Y.; Pilch, P.F.; Bell, G.I.; Seino, S.
J. Biol. Chem. 264, 7776-7779, 1989
A;Title: Cloning and characterization of the major insulin-responsive glucose transporter
A;Reference number: A33801; MUID:89255193; PMID:2656669
A;Accession: A33801
A;Molecule type: mRNA
A;Residues: 1-509 <FUK>
A;Cross-references: GB:M20747; NID:9186552; PIDN:AAAS9189.1; PID:9307076
R;Buse, J.B.; Yasuda, K.; Lay, T.P.; Seo, T.S.; Liu, M.L.; Olson, A.L.; Pessin, J.E.; Mc
J. Biol. Chem. 267, 11673-11676, 1992
A;Title: Expression and regulation of the human GLUT4/muscle-fat facilitative glucose tr
A;Reference number: I55358; MUID:92291025; PMID:1601840
A;Accession: I55358
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: DNA
A;Residues: 1-151,156-509 <RES>
A;Cross-references: GB:M91463; NID:9183295; PIDN:AAAS2569.1; PID:g183296
R;Chiaramonte, R.; Martini, R.; Taramelli, R.; Comi, P.
Gene 130, 307-308, 1993
A;Title: Identification of the 5' end of the gene encoding a human insulin-responsive gl
A;Reference number: I37425; MUID:93366193; PMID:7516714
A;Accession: I37425
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-11 <RE2>
A;Cross-references: EMBL:X58489; NID:9402583; PIDN:CAA41399.1; PID:g402584
C;Genetics:
A;Gene: GDB:SLC2A4; GLUT4
A;Cross-references: GDB:119997; OMIM:138190
A;Map position: 17p13-17p13
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein
Query Match 30.1%; Score 874; DB 2; Length 509;
Best Local Similarity 38.7%; Pred. No. 1e-61;
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;
QY 57 LVASLAGFSGFLYGNLSVNPAPPIKAFYNESWERRHG----REIDPDTLTWSV 112
DB 24 LVAVFSVAVLSLQFGNIGVFNARQVKVTEQSYNTWLGROGPEGPSIPPCTTULWAL 83
QY 113 TVSIFAGIGVGLTIIVKMKVGLGRKHTLLANNFPAISAAALIMACSLQAGAFEMLIIGRF 172
DB 84 SVAIFSVGWISFLIGLISQWLGKRAMLVNNVLAVLGSLMGLANAAASYEMLIIGRF 143

QY 173 IMGIDGGVALSVLPMYLSSEIPKIRSGLOVTAIFICIGVFTQQLIGLPELICKSSTWP 232
DB 144 LIGAYSGLTSPVPMYGEIAPTHLRGALGTNLQIAIVIGILIAQVIGLSLGLTASLWP 203
QY 233 YLFGVIVPVPVAVVQVLLSPDPSRYLLEKHENARAVKAFQTELPGLKADVSQVEEVLAE 292
DB 204 LLGUTVLPALLQLVLLFPCEPSPRYLYIQLNLPARKSLKRLTGWADVSGVLAELKDE 263
QY 293 SHVQSRILSVLELLRAPYVYRWQVTVVITMACYQLCGLNAILWFTYNSIFGKAGI--PPA 351
DB 264 KRKLREPLSLQLLQSGRTHRQPLIIAVVLQSQLSGINAVFYYSIFETAGVQCPA 323
QY 352 KIPYVTLSTGGIETLAAVFSGLVIEHLGRRLP-LIGFGMLGLFFGFTLTTLTIQDHAPW 410
DB 324 ---YATIGAGVVNTVFTLVSVLLVERAGRTLHLGLAGMCGCAI-LMTVALLLERVPA 379
QY 411 VYLSIVGLIATIASFCGPGGIPFILTGEFFQSQRPAAPIIAGTWNLSNFAVGLLFP 470
DB 380 MGYVSIIVAFGEVAFETGPGPIFWFAELFSQGPAPAAVAGPSNWTNFIIGMGFQ 439
QY 471 FIQKSLDNYCFLVPATICITGAILYXFLVLPETKRTYAEISQAKNKAYPPEEK 526
DB 440 YVAEAMGPVYVLPFALLVGLLGFPIFTFLRVBPETGRITFDQISAAHRTPSLLEQEVK 495
RESULT 12
A32101
glucose transport protein, muscle - rat
N;Alternate names: insulin-responsive glucose transporter
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Aug-1989 #sequence_revision 31-Dec-1993 #text_change 24-Sep-1999
C;Accession: A32101; S03349; A32387; I55419
R;Charon, M.J.; Brosius III, F.C.; Alper, S.L.; Lodish, H.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2535-2539, 1989
A;Title: A glucose transport protein expressed predominantly in insulin-responsive tis
A;Reference number: A32101; MUID:89203263; PMID:2649883
A;Accession: A32101
A;Molecule type: mRNA
A;Residues: 1-509 <CHA>
A;Cross-references: GB:L04524
R;James, D.E.; Strube, M.; Mueckler, M.
Nature 338, 83-87, 1989
A;Title: Molecular cloning and characterization of an insulin-regulatable glucose tran
A;Reference number: S03349; MUID:89143771; PMID:2645527
A;Accession: S03349
A;Molecule type: mRNA
A;Residues: 1-509 <JAM>
A;Cross-references: EMBL:X14771; NID:956501; PIDN:CAA32879.1; PID:g56502
R;Birbaum, M.J.
Cell 57, 305-315, 1989
A;Title: Identification of a novel gene encoding an insulin-responsive glucose transpo
A;Reference number: A32387; MUID:89195241; PMID:2649253
A;Accession: A32387
A;Molecule type: mRNA
A;Residues: 1-348, Q, 350-509 <BIR>
A;Cross-references: GB:M25482; NID:9537942; PIDN:AAA41451.1; PID:g537943
R;Liu, M.L.; Olson, A.L.; Edgington, N.P.; Moye-Rowley, W.S.; Pessin, J.E.
J. Biol. Chem. 269, 28514-28521, 1994
A;Title: Myocyte enhancer factor 2 (MEF2) binding site is essential for C2C12 myotube--
A;Reference number: I55419; MUID:95050643; PMID:7545962
A;Accession: I55419
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-50 <RES>
A;Cross-references: GB:L36125; NID:g536863; PIDN:AAA65751.1; PID:g536864
C;Genetics:
A;Gene: GLUT4
A;Introns: 11/3
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein

Query Match 32.3%, Score 930.57, E 1e-66
 Best Local Similarity 39.6%, Pred. No. 1e-66;
 Matches 188; Conservative 98; Mismatches 178; Indels 11; Gaps 5;

QY 57 LVASLAGAGSSFLYGYNLVWVAPTYIKAFYNESWERRHGRPIDPDTLLTWSVTYI 116
 DB 12 LMLAVGAGVLSQFGYNTGVINAPQVIEEFYNTQWVHYGSIILFTLLTWSLSVAI 71
 QY 117 FAIGLVGTLLVKNMGKVLGRKHTLLANNGFSAISALLMACSLQAGAFEMLVGRIMGI 176
 DB 72 FSVGGMGSPFVGLFVNRFGFRNSMLMNNLLAFVSAVLMGFSKLGKSFEMILIGRIIGV 131
 QY 177 DGGVALSVLPMYLSSEISPKIRSGVTAIFICIGVFTQQLGLPELIGKSTWYPLFG 236
 DB 132 YCGLTTGFPVPMYGEVSPFALRGALGTQLHQLGIWVGILIAQVFGDLSIMGNKDLWPLLS 191
 QY 237 VIVVPAVQVLLSLPFLPDSRYLLLEKHNEARAVKAFQTLGKADYSQVEEVLASHVQ 296
 DB 192 IIFPALQCLILFFCPESPRFLINRNEENRAKSVLKKLRGTADVTHDQENKESRQM 251
 QY 297 RSIRLVSLELLRAPYVWVVTIVTMACYQLCGLNALWYFNSIFGKAGIPPAKIPV 356
 DB 252 MREKVVILELFRSPAYRQPIILIAVLQSLQSGINAVFYSTISIFEXAGV--QQPVA 309
 QY 357 TLSTGGIETLAAVPSGLVIEHLGRPL-LIGGFLMGLFPGTITITLTQDHAPWVYLS 415
 DB 310 TIGSGIVNTAFTVVSLEFVVERAGRRTLHLIGLAGMAGCAI-LMTIALALLEOLPWNMSYLS 368
 QY 416 IVGTLAIASFCSGPGIPILLGEPFQSQORPAAFIAGTVNWLNSNEAVGLLFPPIQKS 475
 DB 369 IVAIFGVAFVFEVGGPIPMFIVAEFSQGPRAAIAVAGFSNWTNFINVGMCFQVQOL 428
 QY 476 LDTYCLVFAFICITGAIYLYFVLPEKNTYAEISQAF-----SKRNKAYPPEE 525
 DB 429 CGPVFIIFTVLLVLPFIYFKVPTKGTDEIASGFRQGGASQSDKT--PEE 481

RESULT 9
 A25949
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 24-Sep-1999
 C:Accession: A25949; I55274
 R:Binbaum, M.J.; Haepel, H.C.; Rosen, O.M.
 A:Title: Cloning and characterization of a cDNA encoding the rat brain glucose-transporter
 A:Reference number: A25949; MUID:86287338; PMID:3016720
 A:Accession: A25949
 A:Molecule type: mRNA
 A:Residues: 1-492 <BIR>
 A:Cross-references: GB:M13979; NID:G204413; PIDN:AAA41248.1; PID:G204414
 A:Experimental source: brain
 R:Williams, S.A.; Binbaum, M.J.
 J. Biol. Chem. 263, 19513-19518, 1988
 A:Title: The rat facilitated glucose transporter gene.
 A:Reference number: I55274; MUID:89066774; PMID:3198639
 A:Accession: I55274
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-492 <RES>
 A:Cross-references: GB:M2063; NID:G204531; PIDN:AAA41297.1; PID:G204533
 A:Experimental source: brain
 C:Genetics:
 A:Introns: 6/3; 38/3; 92/2; 172/3; 227/1; 289/3; 324/3; 358/3; 426/3
 C:Superfamily: glucose transport protein
 C:Keywords: membrane protein

Query Match 32.0%; Score 929.5; DB 2; Length 492;
 Best Local Similarity 39.4%; Pred. No. 3.8e-66;
 Matches 187; Conservative 98; Mismatches 179; Indels 11; Gaps 5;

QY 57 LVASLAGAGSSFLYGYNLVWVAPTYIKAFYNESWERRHGRPIDPDTLLTWSVTYI 116
 DB 12 LMLAVGAGVLSQFGYNTGVINAPQVIEEFYNTQWVHYGSIILFTLLTWSLSVAI 71
 QY 117 FAIGLVGTLLVKNMGKVLGRKHTLLANNGFSAISALLMACSLQAGAFEMLVGRIMGI 176
 DB 72 FSVGGMGSPFVGLFVNRFGFRNSMLMNNLLAFVSAVLMGFSKLGKSFEMILIGRIIGV 131
 QY 177 DGGVALSVLPMYLSSEISPKIRSGVTAIFICIGVFTQQLGLPELIGKSTWYPLFG 236
 DB 132 YCGLTTGFPVPMYGEVSPFALRGALGTQLHQLGIWVGILIAQVFGDLSIMGNKDLWPLLS 191
 QY 237 VIVVPAVQVLLSLPFLPDSRYLLLEKHNEARAVKAFQTLGKADYSQVEEVLASHVQ 296
 DB 192 IIFPALQCLILFFCPESPRFLINRNEENRAKSVLKKLRGTADVTHDQENKESRQM 251
 QY 297 RSIRLVSLELLRAPYVWVVTIVTMACYQLCGLNALWYFNSIFGKAGIPPAKIPV 356
 DB 252 MREKVVILELFRSPAYRQPIILIAVLQSLQSGINAVFYSTISIFEXAGV--QQPVA 309
 QY 357 TLSTGGIETLAAVPSGLVIEHLGRPL-LIGGFLMGLFPGTITITLTQDHAPWVYLS 415
 DB 310 TIGSGIVNTAFTVVSLEFVVERAGRRTLHLIGLAGMAGCAI-LMTIALALLEOLPWNMSYLS 368
 QY 416 IVGTLAIASFCSGPGIPILLGEPFQSQORPAAFIAGTVNWLNSNEAVGLLFPPIQKS 475
 DB 369 IVAIFGVAFVFEVGGPIPMFIVAEFSQGPRAAIAVAGFSNWTNFINVGMCFQVQOL 428
 QY 476 LDTYCLVFAFICITGAIYLYFVLPEKNTYAEISQAF-----SKRNKAYPPEE 525
 DB 429 CGPVFIIFTVLLVLPFIYFKVPTKGTDEIASGFRQGGASQSDKT--PEE 481

QY 117 FAIGLVGTLLVKNMGKVLGRKHTLLANNGFSAISALLMACSLQAGAFEMLVGRIMGI 176
 DB 72 FSVGGMGSPFVGLFVNRFGFRNSMLMNNLLAFVSAVLMGFSKLGKSFEMILIGRIIGV 131
 QY 177 DGGVALSVLPMYLSSEISPKIRSGVTAIFICIGVFTQQLGLPELIGKSTWYPLFG 236
 DB 132 YCGLTTGFPVPMYGEVSPFALRGALGTQLHQLGIWVGILIAQVFGDLSIMGNADLWPLLS 191
 QY 237 VIVVPAVQVLLSLPFLPDSRYLLLEKHNEARAVKAFQTLGKADYSQVEEVLASHVQ 296
 DB 192 VIFPALQCLILFFCPESPRFLINRNEENRAKSVLKKLRGTADVTHDQENKESRQM 251
 QY 297 RSIRLVSLELLRAPYVWVVTIVTMACYQLCGLNALWYFNSIFGKAGIPPAKIPV 356
 DB 252 MREKVVILELFRSPAYRQPIILIAVLQSLQSGINAVFYSTISIFEXAGV--QQPVA 309
 QY 357 TLSTGGIETLAAVPSGLVIEHLGRPL-LIGGFLMGLFPGTITITLTQDHAPWVYLS 415
 DB 310 TIGSGIVNTAFTVVSLEFVVERAGRRTLHLIGLAGMAGCAV-LMTIALALLEOLPWNMSYLS 368
 QY 416 IVGTLAIASFCSGPGIPILLGEPFQSQORPAAFIAGTVNWLNSNEAVGLLFPPIQKS 475
 DB 369 IVAIFGVAFVFEVGGPIPMFIVAEFSQGPRAAIAVAGFSNWTNFINVGMCFQVQOL 428
 QY 476 LDTYCLVFAFICITGAIYLYFVLPEKNTYAEISQAF-----SKRNKAYPPEE 525
 DB 429 CGPVFIIFTVLLVLPFIYFKVPTKGTDEIASGFRQGGASQSDKT--PEE 481

RESULT 10
 S04223
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 24-Sep-1999
 C:Accession: S04223
 R:Weiler-Guettler, H.; Zinke, H.; Moeskel, B.; Frey, A.; Gassen, H.G.
 A:Title: cDNA cloning and sequence analysis of the glucose transporter from porcine bio
 A:Reference number: S04223; MUID:89302689; PMID:2472815
 A:Accession: S04223
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-451 <WEI>
 A:Cross-references: GB:X17058; NID:G1955; PIDN:CAA34904.1; PID:G1956
 C:Superfamily: Glucose transport protein
 C:Keywords: transmembrane protein

Query Match 30.2%; Score 877.5; DB 2; Length 451;
 Best Local Similarity 39.9%; Pred. No. 4.7e-62;
 Matches 177; Conservative 92; Mismatches 164; Indels 11; Gaps 5;

QY 88 FYNESWERRHGRPIDPDTLLTWSVTYI FAFGLVGTLLVKNMGKVLGRKHTLLANNGF 147
 DB 2 FYNQTLHRYGESISPATLLTMSLSVAIFSVMGSGFVFNRFGRNMLMMLL 61
 QY 148 AISAAALMACSLQAGAFEMLVGRFTMGIDGVALSVLPMYLSSEISPKIRSGISQVTAI 207
 DB 62 AFISAVLMGFSKLGKSFEMILIGRIIGVTCGTGTFVPMYGEVSPFALRGALGTQLHQL 121
 QY 208 FICIGVFTQQLGLPELIGKSTWYPLFGVIVVPAVQVLLSLPFLPDSRYLLLEKHNEA 267
 DB 122 GIVVGILIAQVFGDLSIMGNELWPLLSVIFIPALLQCVLLPFCPESPRFLINRNEEN 181
 QY 268 RAVKAFQTLGKADYSQVEEVLASHVQSRSLVLELLRAPYVWVVTIVTMACY 327
 DB 182 RAKSVLKKLRGTADVTHDQENKESRQMRKMKVITILEFRSAAYROPILIAVLQSLQ 241
 QY 328 QLCGLNAILWYFNSIFGKAGIPPAKIPYVTLSTGGIETLAAVPSGLVIEHLGRRL-LIG 386
 DB 242 QLSGINAVFYSTISIFEXAGV--QQPVAITGSGIVNTAFTVVSLEFVVERAGRRTLHLIG 299

Development 116, 555-561, 1992
A>Title: Differential screening of a PCR-generated mouse embryo cDNA library: glucose th
A/Reference number: I48366
A/Accession: I48366
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 152-237 <RE2>
A/Cross-references: EMBL:X69697; NID:g312592; PIDN:CAA49367.1; PID:g312593
C/Genetics:
A/Gene: GLUT1
C/Superfamily: glucose transport protein
C/Keywords: transmembrane protein

Query Match 32.4%; Score 941.5; DB 2; Length 492;
Best Local Similarity 39.8%; Pred. No. 4.2e-67;
Matches 189; Conservative 97; Mismatches 178; Indels 11; Gaps 5;

QY 57 LVASLAGAGSGFLVGNLSVNNATPYIKAFYNESWERRGRIDPDITLLWSVTSI 116
DB 12 LMLAVGAGVLGSLQFGYNTGVNAPQKVFIEFYNTWNRHYGEIPSTTTLLMSLVAI 71
QY 117 FAIGLVGLTLLVIMIGKVLGRKHTLLANNPAISALLMACSLQAGAFEMLIIVGRFMGI 176
DB 72 FSVGGMIGSFGVGLFVNFRGRNSMLMNNLLAFVAVLGMGFSKLAKSFEMLLIGRFIIGV 131
QY 177 DGGVALSVLPMYLSISPKIRSGISQVTAIFICIGVFTGQLLGLPBLGKESWPYLF 236
DB 132 YCGLTTGTFVPMYVGEVSPPTALRGALGTLHQLGIIVGILIAQVFGLDISMGNDLWPLLS 191
QY 237 VIVVPAVQVLLSLPPLDPSRVLLEKNEARAVKAFQTFGLKADYDQVEEVLAEHVQ 296
DB 192 VIFIPALLOCIILLPCPESPRFLINNEENRAKSVLKURGTADVTRDLOQEMKEGRQM 251
QY 297 RSRLVSVLELLRAPYVRQVVTVMACVQLCGINAIWFYTNISIFGKAGIPPAKIPYV 356
DB 252 MREKVTIILELFRSPAYRQPIIAVVLQSQLSGINAVFYYSISEKAGV--QQPVYA 309
QY 357 TLSTGGIETLAAVFSGVLVIEHGRPL-LIGFGMLGFFFTGLTITLTDHAPWPVYLS 415
DB 310 TIGSGIVNTAFTVSLFVVERAGRRTLHLGLAGMACAVLMTIALALLERLPWMSYLS 368
QY 416 IGVILAIASFGSGGPIPIITGTFEQSQORPAAFIAGTVNLSNFAVGLLFPFFIQKS 475
DB 369 IVAIFGVFAFFVEFGPIPIVFAELFSGQPRPAIAVAGFSNWTNFIIVGCMFQYVEQL 428
QY 476 LDTYCLVFPATICITGAIYLYFVLPEKTRNYVAEISQAF-----SKNKAYPPPE 525
DB 429 CGPYVFIITVLLVFFIYFKVPETKGRTFDEIASGFRQGGASQSDKT--PEE 481

RESULT 7
A30797
glucose transport protein - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 01-Dec-1989 #sequence_revision 30-Sep-1991 #text_change 24-Sep-1999
C/Accession: A30797
R/Asano, T.; Shibasaki, Y.; Kasuga, M.; Kanazawa, Y.; Takaku, F.; Akanuma, Y.; Oka, Y.
Biochem. Biophys. Res. Commun. 154, 1204-1211, 1988
A/Title: Cloning of a rabbit brain glucose transporter cDNA and alteration of glucose tr
A/Reference number: A30797; MUID:88309104; PMID:3408493
A/Accession: A30797
A/Molecule type: mRNA
A/Residues: 1-492 <ASA>
A/Cross-references: EMBL:M21747; NID:g165633; PIDN:AAA1444.1; PID:g165634
C/Superfamily: glucose transport protein
C/Keywords: transmembrane protein

Query Match 32.4%; Score 940.5; DB 2; Length 492;
Best Local Similarity 39.7%; Pred. No. 5e-67;
Matches 188; Conservative 98; Mismatches 179; Indels 9; Gaps 3;

QY 57 LVASLAGAGSGFLVGNLSVNNATPYIKAFYNESWERRGRIDPDITLLWSVTSI 116

12 LMLAVGAGVLGSLQFGYNTGVNAPQKVFIEFYNTWNRHYGEIPSTTTLLMSLVAI 71
117 FAIGLVGLTLLVIMIGKVLGRKHTLLANNPAISALLMACSLQAGAFEMLIIVGRFMGI 176
72 FSVGGMIGSFGVGLFVNFRGRNSMLMNNLLAFVAVLGMGFSKLAKSFEMLLIGRFIIGV 131
177 DGGVALSVLPMYLSISPKIRSGISQVTAIFICIGVFTGQLLGLPBLGKESWPYLF 236
132 YCGLTTGTFVPMYVGEVSPPTALRGALGTLHQLGIIVGILIAQVFGLDISMGNDLWPLLS 191
237 VIVVPAVQVLLSLPPLDPSRVLLEKNEARAVKAFQTFGLKADYDQVEEVLAEHVQ 296
192 VIFIPALLOCIILLPCPESPRFLINNEENRAKSVLKURGTADVTRDLOQEMKEGRQM 251
297 RSRLVSVLELLRAPYVRQVVTVMACVQLCGINAIWFYTNISIFGKAGIPPAKIPYV 356
252 MREKVTIILELFRSPAYRQPIIAVVLQSQLSGINAVFYYSISEKAGV--QQPVYA 309
357 TLSTGGIETLAAVFSGVLVIEHGRPL-LIGFGMLGFFFTGLTITLTDHAPWPVYLS 415
310 TIGSGIVNTAFTVSLFVVERAGRRTLHLGLAGMACAVLMTIALALLERLPWMSYLS 368
416 IGVILAIASFGSGGPIPIITGTFEQSQORPAAFIAGTVNLSNFAVGLLFPFFIQKS 475
369 IVAIFGVFAFFVEFGPIPIVFAELFSGQPRPAIAVAGFSNWTNFIIVGCMFQYVEQL 428
476 LDTYCLVFPATICITGAIYLYFVLPEKTRNYVAEISQAF-----SKNKAYPPPE 525
429 CGPYVFIITVLLVFFIYFKVPETKGRTFDEIASGFRQGGASQSDKT--PEE 481

RESULT 8
A27217
glucose transport protein - human
N/Alternate names: monosaccharide transport protein
C/Species: Homo sapiens (man)
C/Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 24-Sep-1999
C/Accession: A27217; A45957; S13065
R/Mueckler, M.; Caruso, C.; Baldwin, S.A.; Panico, M.; Blench, I.; Morris, H.R.; Allar
Science 229, 941-945, 1985
A/Title: Sequence and structure of a human glucose transporter.
A/Reference number: A27217; MUID:85272595; PMID:3839598
A/Accession: A27217
A/Molecule type: mRNA
A/Residues: 1-492 <WUE>
A/Cross-references: GB:K03195; NID:g183302; PIDN:AAA52571.1; PID:g183303
A/Note: Parts of this sequence were confirmed by peptide sequencing
A/Note: This protein was shown not to be processed at amino or carboxyl ends
R/Fukumoto, H.; Seino, S.; Imura, H.; Seino, Y.; Bell, G.I.
Diabetes 37, 657-661, 1988
A/Title: Characterization and expression of human HepG2/erythrocyte glucose-transporte
A/Reference number: A45957; MUID:88196648; PMID:2834252
A/Accession: A45957
A/Molecule type: DNA
A/Residues: 1-6 <FUK>
A/Cross-references: GB:M20653; NID:g183974; PIDN:AAB61084.1; PID:g463112
R/Wadzinski, B.E.; Shanahan, M.F.; Seamon, K.B.; Ruoho, A.E.
Biochem. J. 272, 151-158, 1990
A/Title: Localization of the forskolin photolabelling site within the monosaccharide t
A/Reference number: S13065; MUID:91090708; PMID:2264820
A/Accession: S13065
A/Status: preliminary
A/Molecule type: protein
A/Residues: 254-268 <WAD>
C/Genetics:
A/Gene: GDB:SLC2A1; GLUT; GLUT1
A/Cross-references: GDB:120627; OMIM:138140
A/Map position: 1p35-p31.3
C/Superfamily: glucose transport protein
C/Keywords: glycoprotein; transmembrane protein
F/45/Binding site: carbohydrate (Asn) (covalent) #status experimental
F/411/Binding site: carbohydrate (Asn) (covalent) #status absent

I45902
glucose transporter type I - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 24-Sep-1999
C/Accession: I45902
R/Boado, R.J.; Pardridge, W.M.
Mol. Cell. Neurosci. 1, 224-232, 1991
A/Title: Molecular cloning of the bovine blood-brain barrier glucose transporter cDNA and its complementary DNA
A/Reference number: I45902
A/Accession: I45902
A/Status: preliminary; translated from GB/ENBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-492 <BOA>
A/Cross-references: GB:M60448; NID:G163104; PIDN:AAA30550.1; PID:G163105
C/Genetics:
A/Genes: GLUT-1
C/Superfamily: glucose transport protein

Query Match 32.7%; Score 949.5; DB 2; Length 492;
Best Local Similarity 40.0%; Pred. No. 9.7e-68;
Matches 190; Conservative 100; Mismatches 174; Indels 11; Gaps 5;

Qy 57 LVASLAGAFSSFLYGYNLVSVNAPTYIKAFYNESWERRHGRPDPTLTLLWSVTYSI 116
Db 12 LMLAVGAGVLSLQFGVNTGVINAPQKVIIEFNQTVQYGEPIPPATLTLLWSLVAI 71
Qy 117 FATGGLVGLTVIKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLVGRFIMI 176
Db 72 FSVGGMIGSPVGLFVNRFGRRNSMLMMLLAFVSAVLMGFSKLGKSPFEMLTIGRFIIGV 131
Qy 177 DGVVALSVLPMYLSSETSPKEIRSLQVTAIFICIGVFTGQLGLPELLGKSTWPFYLF 236
Db 132 YCLTTFGVPMYGVESPTLGRALGTLHQLGVVGLIAGVGLDSINGQELMPLLS 191
Qy 237 VIIVPAVQVLLSPFLPDSPRYLLEKHEARAVKAFQFFLGKADVSQVEVEVLAESHVQ 296
Db 192 VIFIPALLQCLLPFCPSPRFLINNEENAKSVLKLRTADVTDRDQEMKEESQOM 251
Qy 297 RSRLVSVLELLRAPYVRVQVTVIVTMACYLQGLNAIWFYTNIFPKAGIPPAKIPYV 356
Db 252 MREKKVTILELPSAAVROPILLIAVVLQSLSGINAVFYSTSIPEKAGV--QQPYA 309
Qy 357 TLSTGGIETLAAVFGSLVTEHLGRPL-LIGFGFLMGLFPFGTLTTLTLODHAPWPYLS 415
Db 310 TIGSGIVNTAFTVSVLFVVERAGRTLHLIGLAGMAGCAV-LMTTALLERLPMWSYLS 368
Qy 416 IVGILAITASFCSGPGGIPFILTGFQSQRPAAFIAGTVNWSNFAVGLLRPFIOKS 475
Db 369 IVAIFGVFAFFVGVGPDPFWFIVAEFLPSQGRPAIAVAGFSNWTNFIWGMCFQYVEQL 428
Qy 476 LDTYCLFLVATICITGAIYLYFVLPETKRTYAEISQAF-----SKRNKAYPPEE 525
Db 429 CGSYVFIETVLLVLFIFTYFKVPETKRTFDEIASGPRQGGASQSDKT--PEE 481

RESULT 5
A41264
glucose transport protein 3 - chicken
C/Species: Gallus gallus (chicken)
C/Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 04-Sep-1998
C/Accession: A41264
R/White, M.K.; Rall, T.B.; Weber, M.J.
Mol. Cell. Biol. 11, 4448-4454, 1991
A/Title: Differential regulation of glucose transporter isoforms by the src oncogene in chicken erythrocytes
A/Reference number: A41264; MUID:91342646; PMID:1875932
A/Accession: A41264
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-496 <WHI>
A/Cross-references: GB:M37785
C/Superfamily: glucose transport protein
C/Keywords: transmembrane protein

Query Match 32.5%; Score 942.5; DB 2; Length 496;
Best Local Similarity 39.3%; Pred. No. 3.5e-67;
Matches 193; Conservative 94; Mismatches 187; Indels 17; Gaps 5;

Qy 49 RKMDSLLVASLAGAFSSFLYGYNLVSVNAPTYIKAFYNESWERRHGRPDPTLTLL 108
Db 4 KKKITASLYAVSVAAG-SLQFGVNTGVINAPKIIQAFYNTUSQSRGETISPELLTS 62
Qy 109 LWSVTYSIFAIGSLVGLTVIKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLI 168
Db 63 LWSLSVAIFSVGGMIGSPVGLFVNRFGRRNSMLVNLVAFAGGALMALSKIKAEMLI 122
Qy 169 VGFPMINGDGVVALSVLPMYLSSEISPKIRSLQVTAIFICIGVFTGQLGLPELLGKE 228
Db 123 IGRFIIIGLFCGLCTGFPVPMYIIEVSPSTSLRGAFGLNQLGIVVGLVAQIFGLEGIMGT 182
Qy 229 STWPYLFVGVVPAVQVLLSPFLPDSPRYLLEKHEARAVKAPQTFGLKADVSQVEVE 288
Db 183 ALWPLLGLFTIIPAVLQCVALLFCPSPRFLINKEBEKAQTVLQKLRGTQDVSQDISE 242
Qy 289 VLAESHVQSRISRLSVLELLRAPYVRVQVTVIVTMACYLQGLNAIWFYTNISFGKAGI 348
Db 243 MKESAKSMEKQKATVLELFRSPVNRQPIIISITLQSLQSLSGINAVFYSTGTGIFERAGI 302
Qy 349 PPAKIPYVTLSTGGIETLAAVFGSLVTEHLGRPL-LIGFGFLMGLFPFGTLTTLTLODHA 408
Db 303 --TQPYATIGAGVNTVFTVSVLSFLVERAGRTLHLVGLGMVACAAWTTIALAKE-- 358
Qy 409 PWPYLSIVGILAITASFCSGPGGIPFILTGFQSQRPAAFIAGTVNWSNFAVGLL 468
Db 359 KMRYSIVATGCVFALFEIGPDPFWFIVAEFLPSQGRPAIAVAGFSNWTNFIWGLM 418
Qy 469 FPIQSLDTCFLVATICITGAIYLYFVLPETKRTYAEISQAFSKRNKAYPPEEKID 528
Db 419 FPVAKLCPYVFLIFLPLVFLIFLFFITFTYFKVPETKRTFEDISRGF-----EEQVE 469
Qy 529 SAVTDAPASSP 539
Db 470 ---TSPSPSP 477

RESULT 6
S09705
glucose transport protein GT1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 17-Nov-2000
C/Accession: S09705; A30310; I49697; I48366
R/Reed, B.C.; Shade, D.; Alperovich, F.; Vang, M.
Arch. Biochem. Biophys. 279, 261-274, 1990
A/Title: 3T3-L1 adipocyte glucose transporter (HepG2 class): sequence and regulation of expression in 3T3-L1 adipocytes
A/Reference number: S09705; MUID:90274408; PMID:2190533
A/Accession: S09705
A/Molecule type: mRNA
A/Residues: 1-492 <REE>
A/Cross-references: EMBL:M22998; NID:G193551; PIDN:AAA37707.1; PID:G309262
R/Kaestner, K.H.; Christy, R.J.; McLenithan, J.C.; Braiterman, L.T.; Cornelius, P.; Pek
Proc. Natl. Acad. Sci. U.S.A. 86, 3150-3154, 1989
A/Title: Sequence, tissue distribution, and differential expression of mRNA for a putative glucose transporter in rat liver and adipose tissue
A/Reference number: A30310; MUID:89240694; PMID:2654938
A/Accession: A30310
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-51, 153-192, 194, 196-402, 404-464, 466-492 <KAE>
R/Murakami, T.; Nishiyama, T.; Shirotani, T.; Shinohara, Y.; Kan, M.; Ishii, K.; Kanai, Y.
J. Biol. Chem. 267, 9300-9306, 1992
A/Title: Identification of two enhancer elements in the gene encoding the type 1 glucosyltransferase
A/Reference number: I49697; MUID:92250534; PMID:1339457
A/Accession: I49697
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-6 <RES>
A/Cross-references: GB:D10229; NID:G220414; PIDN:BAA01081.1; PID:G3582350
R/Smith, D.E.; Gridley, T.

Db 197 GLTGVPAALQLLLPFPESPRLYLKQKDEAAAKALQTLRGWDSVDREVAIRQDEA 256
QY 296 QRSIRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINVYTNISFGKAGIPPAKIPY 355
Db 257 EKAAGFISVVKLFMRSLRWQLLSITVLGGQQLSGVNAIYYADQIYLSAGVPEEHVQY 316
QY 356 VTLSTGGIB---TLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWP 412
Db 317 VTAGTGANVWTFCAVFP---VVELLGRRLLLGLGFSICLIACCVTTLAALQDTVSWMP 373
QY 413 YLSIVGILAIASFCSPGGIPFILTGEFFQOSQORPAAFIAGTVNWLNSFVAGLLPFI 472
Db 374 YISIVCVISVIGHALGPSIPALLITEIFLQSSRPSAFMVGSGVHMLSNFTVGLRPFI 433
QY 473 QKSLDTCFLVPATICITGAIYLVFLPETKRTVABISQAFSKKNK---AYPPEEKI 527
Db 434 QEGLPSPFIVFAVICLLTIIYFVLVPEKTAKTFEINQITKKNKVSSEVYPEKEEL 491

RESULT 2
I53268
glut 5 protein - rat
N;Alternate names: fructose transporter
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1999
A;Accession: I53268; I51895
R;Inukai, K.; Asano, T.; Katagiri, H.; Ishihara, H.; Anai, M.; Fukushima, Y.; Tsukuda, K.
Endocrinology 133, 2009-2014, 1993
A;Title: Cloning and increased expression with fructose feeding of rat jejunal GLUTs.
A;Reference number: I53268; PMID:94008761; PMID:8404647
A;Accession: I53268
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <RES>
A;Cross-references: GB:D13871; NID:9436578; PIDN:BA002983.1; PID:9436579
R;Rand, E.B.; Depauli, A.M.; Davidson, N.O.; Bell, G.I.; Burant, C.F.
Am. J. Physiol. 264, G1169-G1176, 1993
A;Title: Sequence, tissue distribution, and functional characterization of the rat fructose transporter.
A;Reference number: I51895; PMID:93325725; PMID:8333543
A;Accession: I51895
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-163, 'A', 165-284, 'A', 286-451, 'T', 453-478, 'N', 480-502 <RE2>
A;Cross-references: GB:I05195; NID:9204415; PIDN:AAA02627.1; PID:9204416
C;Genetics:
A;Gene: glut 5
C;Superfamily: glucose transport protein

Query Match 38.2%; Score 1109.5; DB 2; Length 502;
Best Local Similarity 44.7%; Pred. No. 1.8e-80;
Matches 217; Conservative 97; Mismatches 167; Indels 5; Gaps 2;
QY 56 LLVASIAGAFSGSFLYGNLSVNVNAPTPIYKAFYNESWERRHGRPIDPDTLLWSVTVS 115
Db 16 LALATPLAAGSSFOQYGVNVAVNSPBFMOQYNDTYDRNKENTIESFTLLWSLTVS 75
QY 116 IFAGLGLVGLTKVMIGKVLGRKHTLLANNNGPAISALLMACSLQAGAFEMLIYGRFTMG 175
Db 76 MPFFGFGISLMVGLVNNLGRKALLFNFIIPAILMGCSTAKSFETIIASRLLVG 135
QY 176 IDGGVALSVLPMLSEISPKIRSGISLQVTAIFICIGVFTGQLLGLPELLGKSTWPLYF 235
Db 136 ICAGISSNVVPMYLGELAPKNLKGALGVVPOLFIVTGLVLAQLGLRSVLASEGWPILL 195
QY 236 GVIVPAAVQLLSLFLPDSPPRYLLEKHNEARAVKAFQTLGKADVSQVEVLAESHV 295
Db 196 GTGVPAGLQLLLPFPESPRLYLKQKDEAAAKALQTLRGWDSVDREVAIRQDEA 255
QY 296 QRSIRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINVYTNISFGKAGIPPAKIPY 355
Db 256 EKAAGFISVVKLFMRQSLRWQLLSITVLGGQQLSGVNAIYYADQIYLSAGVKSNDVQY 315
QY 356 VTLSTGGIETLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWPYLS 415

Db 316 VTAGTGANVWTFWTVFVVELWGRNLLIGFSTCLTACIVLTVALQNTISWMEYVS 375
QY 416 IYVGIILAIASFCSPGGIPFILTGEFFQOSQORPAAFIAGTVNWLNSFVAGLLPFIQKS 475
Db 376 IYCVIVVIGHAVGPSIPALPITEIFLQSSRPSAYMIGSVHMLSNFTVGLIPIFPFQVQ 435
QY 476 LDTYCFVLPATICITGAIYLVFLPETKRTVABISQAFSKKNK---AYPPEEKIDSAVT 532
Db 436 LGPYFIIFAILCLLTIIYFMVWVPEKRTFVINOIFAKKNKVDVYP---EKEEKELN 493
QY 533 DAPASS 538
Db 494 DLPAT 499

RESULT 3
G02864
fructose transporter - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C;Accession: G02864
R;Takeuchi, J.
submitted to the EMBL Data Library, January 1994
A;Reference number: G07991
A;Accession: G02864
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-481 <TAX>
A;Cross-references: EMBL:U11843; NID:9516513; PIDN:AAB0641.1; PID:9516515
C;Genetics:
A;Gene: GLUT5
A;Introns: 11/2; 24/3; 78/2; 120/1; 171/1; 213/1; 275/3; 312/3; 346/3; 372/1; 414/3
C;Superfamily: glucose transport protein

Query Match 35.6%; Score 1032.5; DB 2; Length 481;
Best Local Similarity 44.0%; Pred. No. 2.3e-74;
Matches 209; Conservative 94; Mismatches 155; Indels 17; Gaps 5;
QY 60 SIAGAFSGSFLYGNLSVNVNAPTPIYKAFYNESWERRHGRPIDPDTLLWSVTVSIPA 118
Db 7 SNKEGWTWL-----LSTPQHMLMQOYFNYTGTGTFMEDFDTLLWSVTVSMFP 59
QY 119 IGLVGLTGLVTKVMIGKVLGRKHTLLANNNGPAISALLMACSLQAGAFEMLIYGRFTMG 178
Db 60 FCGFISLVLGVLVGNKFGKALLFNFIIPAILMGCSTAKSFETIIISRLVGLICA 119
QY 179 GVALSVLPMLSEISPKIRSGISLQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFPGVI 238
Db 120 GVSSNVVPMYLGELAPKNLKGALGVVPOLFIVTGLVLAQLGLRNLNLANVDGWPIILLGLT 179
QY 239 VVPAAVQLLSLFLPDSPPRYLLEKHNEARAVKAFQTLGKADVSQVEVLAESHVQVS 298
Db 180 GVGAAQLQLLLPFPESPRLYLKQKDEAAAKALQTLRGWDSVDREVAIRQDEA 239
QY 299 IRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINVYTNISFGKAGIPPAKIPVVT 358
Db 240 AGFISVVKLFMRSLRWQLLSITVLGGQQLSGVNAIYYADQIYLSAGVPEEHVQVTA 299
QY 359 STGGIETLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWPYLS 415
Db 300 GTGAVNVWTFCAVFP---VVELLGRRLLLGLGFSICLIACCVTTLAALQDTVSWMEYIS 356
QY 416 IYVGIILAIASFCSPGGIPFILTGEFFQOSQORPAAFIAGTVNWLNSFVAGLLPFIQKS 475
Db 357 IYCVIVVIGHALGPSIPALLITEIFLQSSRPSAFMVGSGVHMLSNFTVGLIPIFPFQVQ 416
QY 476 LDTYCFVLPATICITGAIYLVFLPETKRTVABISQAFSKKNK---AYPPEEKI 527
Db 417 LGPYFIIFAILCLLTIIYFVLVPEKTAKTFEINQITKKNKVSSEVYPEKEEL 471

RESULT 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: April 6, 2004, 10:13:22 ; Search time 21 Seconds
(without alignment)
2578.849 Million cell updates/sec
Title: US-09-981-947B-2
Perfect score: 2903
Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWQAAATTATKKEHPL 563
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283366 seqs, 96191536 residues
Total number of hits satisfying chosen parameters: 283366
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 78:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

ALIGNMENTS

RESULT 1

A36629
Glucose transport protein 5 - human
N:Alternate names: fructose transporter
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 17-Mar-1999
C:Accession: A36629; S46486
R:Kayano, T.; Burant, C.F.; Fukumoto, H.; Gould, G.W.; Fan, Y.; Eddy, R.L.; Byers, M.G.
J. Biol. Chem. 265, 13276-13282, 1990
A:Title: Human facilitative glucose transporters. Isolation, functional characterization
adipose tissue and an unusual glucose transporter pseudogene-like sequence (GLUT6).
A:Reference number: A36629; MUID:90330675; PMID:1695905
A:Accession: A36629
A:Molecule type: mRNA
A:Residues: 1-501 <KAY>
A:Cross-references: GB:J05462
R:Mahrach, L.; Takeda, J.; Mesonero, J.; Chantret, I.; Dussaulx, E.; Bell, G.I.; Brot-
Biochem. J. 301, 169-175, 1994
A:Title: Regulation of expression of the human fructose transporter (GLUTs) by cyclic A
A:Reference number: S46486; MUID:94311827; PMID:8037665
A:Accession: S46486
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-11 <MAH>
A:Cross-references: EMBL:U05344
A:Note: the sequence of residues 7-11 and the corresponding nucleotide sequence are not
C:Genetics:
A:Gene: GDB:SLC2A5; GLUT5
A:Cross-references: GDB:119998; OMIM:138230
A:Map position: lp31-lp31
C:Superfamily: glucose transport protein
C:Keywords: glycoprotein; sugar transport; transmembrane protein
F:51/Binding site: carbohydrate (Asn) (Covalent) #status predicted

Result No.	Score	Match	Length	DB	ID	Description
1	1122.5	38.7	501	2	A36629	glucose transport
2	1109.5	38.2	502	2	I53268	glut 5 protein - r
3	1032.5	35.6	481	2	G02864	fructose transport
4	949.5	32.7	492	2	I45902	glucose transport
5	942.5	32.5	496	2	A41264	glucose transport
6	941.5	32.4	492	2	S09705	glucose transport
7	940.5	32.4	492	2	A30797	glucose transport
8	936.5	32.3	492	2	A27217	glucose transport
9	929.5	32.0	492	2	A25949	glucose transport
10	877.5	30.2	451	2	S04223	glucose transport
11	874	30.1	509	2	A33801	muscle-fat glucose
12	854	29.4	509	2	A32101	glucose transport
13	844.5	29.1	493	2	A41751	glucose transport
14	839.5	28.9	508	2	B03110	glucose transport
15	838	28.9	487	2	S53322	fructose transport
16	830	28.6	496	2	A31986	glucose transport
17	816.5	28.1	493	2	S38981	glucose transport
18	772	26.6	522	2	A31556	glucose transport
19	761.5	26.2	523	2	S06920	glucose transport
20	750.5	25.9	533	2	S42330	facilitative gluco
21	721	24.8	528	2	S24344	glucose transport
22	695.5	24.0	524	2	A31318	glucose transport
23	672.5	23.2	576	2	C88950	protein R09B5.11
24	654	22.5	489	2	B53153	glucose transport
25	646	22.3	521	2	A53153	glucose transport
26	582	20.0	505	2	C53153	sugar transporter-
27	511	17.6	560	2	T51485	hypothetical prote
28	505	17.4	569	2	T34295	hypothetical prote
29	480	16.5	488	2	S46118	probable glucose t

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

probable sugar tra
hypothetical prote
metabolite transpo
protein FLN21.12
probable sugar tra
hypothetical prote
probable sugar tra
probable sugar tra
hypothetical prote
D-xylulose-proton sy
L-arabinose transp
hypothetical prote
hypothetical prote

APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 308
LENGTH: 501
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-308

Query Match 38.7%; Score 1122.5; DB 15; Length 501;
Best Local Similarity 46.4%; Pred. No. 8.3e-96;
Matches 222; Conservative 92; Mismatches 155; Indels 9; Gaps 3;

QY 56 LVASLAGAGSSFLYGYNLSVNVNAPTYKAFYNESWERRHGRFIDPDTTLMSVTVS 115
DB 17 LALATLIAAFGSSFGYGVNVAVNSPALLMQQFYNETYGRTEFMEFPLTLMSVTVS 76

QY 116 IFAIGLVGTLIVKMGKVLGRKHTLLANNNGFAISAAALMACSLQAGAFEMLIIVGRFIMG 175
DB 77 MFPGGFTGSLVGLVGNKFGKRGKALLFNNEISIVPALMGCSRVATSFELIIISRLVVG 136

QY 176 IDGGVALSVLPMYLSKPEIRGSIQVTAIFICIGVFTGQLGLPELLGKESWPLYLF 235
DB 137 ICAGVSSNVVPMYLGELAPKNRGLAGVVPQFIVTIGILVAQIFGRNLLANVDGNPILL 196

QY 236 GVIVVPVAVVQLSLPFLPDSPPYLLLEKHEARAVKAFQTFILGKADVSOEVEEVLAEHV 295
DB 197 GLTGVPALQILLPFFESPRLIIQKDEAAKAKALQTLRGWSDVDEVAIEIQEDEA 256

QY 296 QRSIRLVSVLELLRAPYVWQVTVVTVMACYQLCGLNAINWFTYNSIFGKAGIPPAKIPY 355
DB 257 ERAAGFISVLKFRMSLRWQLLSIIVLMGGQQLSGVNAIYYADQIYLSAGVPEEHVQY 316

QY 356 VTLSGGBE---TLAAVFSGLVIEHLGRPLIIGFGLMGLFFGTLITLTLODHAPWVP 412
DB 317 VTAGTGAVNVMTFCVAF---VVELLGRLLLLGFSICLIACCCLTAALALQDVTWMP 373

QY 413 YLSIVGILAIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNWLNSNPAVGLLFFPI 472
DB 374 YISIVCVISYVIGHALGPSPIFALLITEIFLOSSRPSAPWCGSVHLSNFTVGLIFPI 433

QY 473 QKSLDTYCFVLPATICITGAIYLYFVLPETKNTYAEISQAFSKRK---AYPPEEKI 527
DB 434 QEGGLGPSYFIVFAVICLLTTIYIFLIVPETKAKTTEINQIFTKMKVSEVYPEKEEL 491

;; PRIOR APPLICATION NUMBER: 60/360814
;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/303828
;; PRIOR FILING DATE: 2001-09-07
;; PRIOR APPLICATION NUMBER: 60/323380
;; PRIOR FILING DATE: 2001-09-19
;; PRIOR APPLICATION NUMBER: 60/361133
;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/304016
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: 60/304502
;; PRIOR FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 60/305262
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/373881
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: 60/305673
;; PRIOR FILING DATE: 2001-07-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 368
;; SOFTWARE: Custom
;; SEQ ID NO 88
;; LENGTH: 396
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-188-186-88

Query Match 65.8%; Score 1911; DB 12; Length 396;
Best Local Similarity 99.5%; Pred. No. 1.3e-169;
Matches 376; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 137 RKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 196
DB 11 RKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 70

QY 137 IRGSLGQVTAIFCIGVFTGQLLGPPELLGKSTWYPLFGVIVVPAVVQLLSLPFLDPS 256
DB 71 IRGSLGQVTAIFCIGVFTGQLLGPPELLGKSTWYPLFGVIVVPAVVQLLSLPFLDPS 130

QY 257 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 316
DB 131 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 190

QY 317 VVTIVTMACYQLCGLNALWFYNTSIFGKAGIPPAKIPVVTLSGTGIELAAVFSGLVIE 376
DB 191 VVTIVTMACYQLCGLNALWFYNTSIFGKAGIPPAKIPVVTLSGTGIELAAVFSGLVIE 250

QY 377 HLGRRPLLIIGGFLMGLFFGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 436
DB 251 HLGRRPLLIIGGFLMGLFFGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 310

QY 437 LTGEFFQOSORPAAFIIAGTVNWLNSFVAVGLLFPFIQKSLDTCYFLVPATICITGAIYLY 496
DB 311 LTGEFFQOSORPAAFIIAGTVNWLNSFVAVGLLFPFIQKSLDTCYFLVPATICITGAIYLY 370

QY 497 FVLPTKNTYAEISQAF 514
DB 371 FVLPTKNTYAEISQAF 388

RESULT 14
US-10-188-186-82
; Sequence 82, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814

;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/303828
;; PRIOR FILING DATE: 2001-09-07
;; PRIOR APPLICATION NUMBER: 60/323380
;; PRIOR FILING DATE: 2001-09-19
;; PRIOR APPLICATION NUMBER: 60/361133
;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/304016
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: 60/304502
;; PRIOR FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 60/305262
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/373881
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: 60/305673
;; PRIOR FILING DATE: 2001-07-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 368
;; SOFTWARE: Custom
;; SEQ ID NO 82
;; LENGTH: 362
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-188-186-82

Query Match 59.3%; Score 1721; DB 12; Length 362;
Best Local Similarity 90.5%; Pred. No. 6.4e-152;
Matches 342; Conservative 1; Mismatches 1; Indels 34; Gaps 1;

QY 137 RKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 196
DB 11 RKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 70

QY 137 IRGSLGQVTAIFCIGVFTGQLLGPPELLGKSTWYPLFGVIVVPAVVQLLSLPFLDPS 256
DB 71 IRGSLGQVTAIFCIGVFTGQLLGPPELLGKSTWYPLFGVIVVPAVVQLLSLPFLDPS 130

QY 257 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 316
DB 131 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 190

QY 317 VVTIVTMACYQLCGLNALWFYNTSIFGKAGIPPAKIPVVTLSGTGIELAAVFSGLVIE 376
DB 191 VVTIVTMACYQLCGLNALWFYNTSIFGKAGIPPAKIPVVTLSGTGIELAAVFSGLVIE 245

QY 377 HLGRRPLLIIGGFLMGLFFGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 436
DB 246 HLGRRPLLIIGGFLMGLFFGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 276

QY 437 LTGEFFQOSORPAAFIIAGTVNWLNSFVAVGLLFPFIQKSLDTCYFLVPATICITGAIYLY 496
DB 277 LTGEFFQOSORPAAFIIAGTVNWLNSFVAVGLLFPFIQKSLDTCYFLVPATICITGAIYLY 336

QY 497 FVLPTKNTYAEISQAF 514
DB 337 FVLPTKNTYAEISQAF 354

RESULT 15
US-10-295-027-308
; Sequence 308, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezzi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.

; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 86
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-86

Query Match 66.4%; Score 1929; DB 12; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.7e-171; Indels 0; Gaps 0;
Matches 378; Conservative 0; Mismatches 0;

QY 137 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 196
Db 11 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 70
QY 197 IRGSLQVTAIFICIGVFTGQLLGLPELLGKSTWYLVGVVAVVQVLLSLPFLDPS 256
Db 71 IRGSLQVTAIFICIGVFTGQLLGLPELLGKSTWYLVGVVAVVQVLLSLPFLDPS 130
QY 257 RYLLEKHNARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 316
Db 131 RYLLEKHNARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 190
QY 317 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIE 376
Db 191 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIE 250
QY 377 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 436
Db 251 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 310
QY 437 LTGEFFQSQORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFAFATCITGAIYLY 496
Db 311 LTGEFFQSQORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFAFATCITGAIYLY 370
QY 497 FVLPEKNTYAEISQAF 514
Db 371 FVLPEKNTYAEISQAF 388

RESULT 12
US-10-188-186-84
; Sequence 84, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046

; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 84
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-84

Query Match 66.1%; Score 1920; DB 12; Length 396;
Best Local Similarity 99.5%; Pred. No. 1.9e-170;
Matches 376; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 137 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 196
Db 11 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 70
QY 197 IRGSLQVTAIFICIGVFTGQLLGLPELLGKSTWYLVGVVAVVQVLLSLPFLDPS 256
Db 71 IRGSLQVTAIFICIGVFTGQLLGLPELLGKSTWYLVGVVAVVQVLLSLPFLDPS 130
QY 257 RYLLEKHNARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 316
Db 131 RYLLEKHNARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 190
QY 317 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIE 376
Db 191 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIE 250
QY 377 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 436
Db 251 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 310
QY 437 LTGEFFQSQORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFAFATCITGAIYLY 496
Db 311 LTGEFFQSQORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFAFATCITGAIYLY 370
QY 497 FVLPEKNTYAEISQAF 514
Db 371 FVLPEKNTYAEISQAF 388

RESULT 13
US-10-188-186-88
; Sequence 88, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05

; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/322380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 90
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-90

Query Match 69.5%; Score 2017; DB 12; Length 422;
Best Local Similarity 99.5%; Pred No. 1.8e-179;
Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 137 RKHTLANNNGFAISALLMACSLQAGAFEMLIVGRFIMGIDGGVALSVLPWYLSISPKE 196
Db 11 RKHTLANNNGFAISALLMACSLQAGAFEMLIVGRFIMGIDGGVALSVLPWYLSISPKE 70
Qy 197 IRGSLGQVTAIPICIGVTGQLLGLPELLGKSTWYLFVGVVPAVQVLLSLPFLDPS 256
Db 71 IRGSLGQVTAIPICIGVTGQLLGLPELLGKSTWYLFVGVVPAVQVLLSLPFLDPS 130
Qy 257 RYLLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 316
Db 131 RYLLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRWQ 190
Qy 317 VVTVTWACVQVLCGLNAINFWYNSIFGKAGIPPAKIPYVTLSTGGIETLAASFGLVIE 376
Db 191 VVTVTWACVQVLCGLNAINFWYNSIFGKAGIPPAKIPYVTLSTGGIETLAASFGLVIE 250
Qy 377 HLGRRPLIGGFLMGLFFGTITITLTLODHAPVWYLSIVGILAIASFCSGPGGIPFI 436
Db 251 HLGRRPLIGGFLMGLFFGTITITLTLODHAPVWYLSIVGILAIASFCSGPGGIPFI 310
Qy 437 LTGEFFQSORPAAFIIAGTVNWSNFAVGLLFPFIQKSLDTCYFLVPAATICITGAIYLY 496
Db 311 LTGEFFQSORPAAFIIAGTVNWSNFAVGLLFPFIQKSLDTCYFLVPAATICITGAIYLY 370
Qy 497 FVLPEKRTVAEISQAFSKRNKAYPPEKIDSATVD 533
Db 371 FVLPEKRTVAEISQAFSKRNKAYPPEKIDSATVD 407

RESULT 10
US-10-168-651-6
; Sequence 6, Application US/1016851
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.

; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameeni R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,75
; 60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 6618083CD1
US-10-168-651-6

Query Match 66.8%; Score 1939; DB 14; Length 416;
Best Local Similarity 97.9%; Pred. No. 3.4e-172;
Matches 380; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 47 RRRKDWSCSLLVASLAGAFSGSFLYGYNLSVNVNAPPTPKAFYNSWERRHGRPDPTL 106
Db 18 RRRKDWSCSLLVASLAGAFSGSFLYGYNLSVNVNAPPTPKAFYNSWERRHGRPDPTL 77
Qy 107 TLLWSVTVSIFAIGLVGTLIVVMICKVLRKHTLANNNGFAISALLMACSLQAGAFEM 166
Db 78 TLLWSVTVSIFAIGLVGTLIVVMICKVLRKHTLANNNGFAISALLMACSLQAGAFEM 137
Qy 167 LIVGRITMGIDGGVALSVLPWYLSISPKEIRGSLGQVTAIFICIGVFTGQLLGLPELLG 226
Db 138 LIVGRITMGIDGGVALSVLPWYLSISPKEIRGSLGQVTAIFICIGVFTGQLLGLPELLG 197
Qy 227 KESTWYPLFGVIVVPAVQVLLSLPFLDPSRVLLEKNEARAVKAFOTFLGKADVSQEV 286
Db 198 KESTWYPLFGVIVVPAVQVLLSLPFLDPSRVLLEKNEARAVKAFOTFLGKADVSQEV 257
Qy 287 EVVLASHVQSRIRLSVLELLRAPYVRWQVTVITMACYQVLCGLNAINFWYNSIFGKA 346
Db 258 EVVLASHVQSRIRLSVLELLRAPYVRWQVTVITMACYQVLCGLNAINFWYNSIFGKA 317
Qy 347 GIPPAKIPYVTLSTGGIETLAASFGLVIEHLGRRPLIGGFLMGLFFGTITITLTLOD 406
Db 318 GIPPAKIPYVTLSTGGIETLAASFGLVIEHLGRRPLIGGFLMGLFFGTITITLTLOD 377
Qy 407 HAPWPYLSIVGILAIASFCSGPGGIP 434
Db 378 HAPWPYLSIVGILAIASFCSGPGVFP 405

RESULT 11
US-10-188-186-86
; Sequence 86, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02

; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 80
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-80

Query Match
Best Local Similarity 80.6%; Score 2341; DB 12; Length 480;
Matches 467; Conservative 0; Mismatches 6; Indels 60; Gaps 1;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLCDHLRSGVPGRRRDKMSCSLIVAS 60
Dd 1 MARKQNRNSKELGLVPLTDDTSHARPPGGRALLECHLRSGVPGRRRDKMSCSLIVAS 60

QY 61 LAGAFGSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
Dd 61 LAGAFGSPFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120

QY 121 GLVGTLLIVKMIGKVLGRKHTLLANNGFASALLMACSLQAGAFEMLIIVGRFIMGDGV 180
Dd 121 GLVGTLLIVKMIGKVLGRKHTLLANNGFASALLMACSLQAGAFEMLIIVGRFIMGDGV 180

QY 181 ALSVLPWYLSEISPKIRSGLSGOVTAIFICIGVFTGQLLGLPELLGKESTWPLYFGVIV 240
Dd 181 ALSVLPWYLSEISPKIRSGLSGOVTAIFICIGVFTGQLLGLPELLGKESTWPLYFGVIV 240

QY 241 PAVVQLLSLFPDPSRYLLEKHEARAVKAFQTFPLGKADVSQVEEVLAEASHVQRSTR 300
Dd 241 PAVVQLLSLFPDLSRYLLEKHEARAVKAFQTFPLGKADVSQVEEVLAEASHVQRSTR 300

QY 301 LVSVELELRAPYRVQVTVITVMACYQLCGLNAINFWYNSIFGKAGIPPAKIPYVTLST 360
Dd 301 LVSVELELRAPYRVQVTVITVMACYQLCGLNAINFWYNSIFGKAGIPPAKIPYVTLST 360

QY 361 GGIETLAAVPSGLVIEHLGRRLPILGFGMLGFFGTLLITLQDHAPWVPLSIVGIL 420
Dd 361 GGIETLAAVPS----- 371

QY 421 AIIASCSGPGGIPRILTGEFFQOSORPAAFIAGTVNWLSPNFAVGLLFPFIQKSLDITYC 480
Dd 372 -----GIPFILTGEFFQOSORPAAFIAGTVNWLSPNFAVGLLFPFIQKSLDITYC 420

QY 481 FLVPFATICTGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
Dd 421 FLVPFATICTGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 473

RESULT 8
US-10-188-186-78
; Sequence 78, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 78
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-78

Query Match
Best Local Similarity 80.5%; Score 2337.5; DB 12; Length 565;
Matches 474; Conservative 0; Mismatches 4; Indels 135; Gaps 2;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLCDHLRSGVPGRRRDK----- 50
Dd 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLCDHLRSGVPGRRRDKPLRSTSSAA 60

QY 51 ----- 50
Dd 51 ----- 50

QY 61 GSSTTYVASAAKIPIFRNWSLAPEKSSHLRLQLRGLALEVKLISLGSKPOSSISWLV 120
QY 51 -----DWSCLLVASLAGAFSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRP 100
Dd 121 ALBETNEIGRDWSCLLVASLAGAFSSFLYGNLSVNVNAPT----- 163

QY 101 IDPDTLLWSVTVSIFAIGSLVGTLLIVKMIGKVLGRKHTLLANNGFASALLMACSLQ 160
Dd 164 -----HTLLANNGFASALLMACSLQ 185

QY 161 AGAFEMLIIVGRFIMGIDGGVALSVLPMYLSEISPKIRSGLSGOVTAIFICIGVFTGQLIG 220
Dd 186 AGAFEMLIIVGRFIMGIDGGVALSVLPMYLSEISPKIRSGLSGOVTAIFICIGVFTGQLIG 245

QY 221 LPPELLGKESTWPLYFGVIVPAVVQQLLSLFPDPSRYLLEKHEARAVKAFQTFPLGKA 280
Dd 246 LPPELLGKESTWPLYFGVIVPAVVQQLLSLFPDLSRYLLEKHEARAVKAFQTFPLGKA 305

QY 281 DVSQVEEVLAEASHVQRSTRILSVLELRAPYRVQVTVITVMACYQLCGLNAINFWYTN 340
Dd 306 DVSQVEEVLAEASHVQRSTRILSVLELRAPYRVQVTVITVMACYQLCGLNAINFWYTN 365

QY 341 SIFGKAGIPPAKIPYVTLSTGGIETLAAVPSGLVIEHLGRRLPILGFGMLGFFGTLLTI 400
Dd 366 SIFGKAGIPPAKIPYVTLSTGGIETLAAVPSGLVIEHLGRRLPILGFGMLGFFGTLLTI 425

QY 401 TLTLODHAPWVPLSIVGILAIITASFCSGPGGIPFILTGEFFQOSORPAAFIAGTVNWL 460
Dd 426 TLTLODHAPWVPLSIVGILAIITASFCSGPGGIPFILTGEFFQOSORPAAFIAGTVNWL 485

QY 461 SNFAVGLLFPFIQKSLDITYCFLVFATICTGAIYLYFVLPETKNRTYAEISQAFSKRNKA 520
Dd 486 SNFAVGLLFPFIQKSLDITYCFLVFATICTGAIYLYFVLPETKNRTYAEISQAFSKRNKA 545

QY 521 YPPEEKIDSAVTD 533
Dd 546 YPPEEKIDSAVTD 558

RESULT 9
US-10-188-186-90
; Sequence 90, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186

; LENGTH: 472
; TYPE: PRT
; ORGANISM: Human
US-10-281-319-4

Query Match 82.5%; Score 2396; DB 14; Length 472;
Best Local Similarity 99.6%; Pred. No. 8.3e-215;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 55 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDPDTLLLSWTV 114
Db 1 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDPDTLLLSWTV 60

QY 115 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIM 174
Db 61 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIM 120

QY 175 GIDGGVALSVLPYMLSEISPKIRSGLCQVTAIFICIGVFTGQLLGLPELLGKSTWPLYL 234
Db 121 GIDGGVALSVLPYMLSEISPKIRSGLCQVTAIFICIGVFTGQLLGLPELLGKSTWPLYL 180

QY 235 FGVIVPAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQFTLKGADVSQVEEVLAEBSH 294
Db 181 FGVIVPAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQFTLKGADVSQVEEVLAEBSR 240

QY 295 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIFWYTNISFGKAGIPPAKIP 354
Db 241 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIFWYTNISFGKAGIPPAKIP 300

QY 355 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFGMLGFLFTGTTITLTLQDHAPWPVYL 414
Db 301 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFGMLGFLFTGTTITLTLQDHAPWPVYL 360

QY 415 SIVGILAIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQK 474
Db 361 SIVGILAIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQK 420

QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
Db 421 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 472

RESULT 7
US-10-188-186-80
; Sequence 80, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; LENGTH: 472
; TYPE: PRT
; ORGANISM: Human
US-09-822-863-4

Query Match 82.5%; Score 2396; DB 9; Length 472;
Best Local Similarity 99.6%; Pred. No. 8.3e-215;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 55 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDPDTLLLSWTV 114
Db 1 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDPDTLLLSWTV 60

QY 115 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIM 174
Db 61 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIM 120

QY 175 GIDGGVALSVLPYMLSEISPKIRSGLCQVTAIFICIGVFTGQLLGLPELLGKSTWPLYL 234
Db 121 GIDGGVALSVLPYMLSEISPKIRSGLCQVTAIFICIGVFTGQLLGLPELLGKSTWPLYL 180

QY 235 FGVIVPAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQFTLKGADVSQVEEVLAEBSH 294
Db 181 FGVIVPAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQFTLKGADVSQVEEVLAEBSR 240

QY 295 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIFWYTNISFGKAGIPPAKIP 354
Db 241 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIFWYTNISFGKAGIPPAKIP 300

QY 355 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFGMLGFLFTGTTITLTLQDHAPWPVYL 414
Db 301 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFGMLGFLFTGTTITLTLQDHAPWPVYL 360

QY 415 SIVGILAIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQK 474
Db 361 SIVGILAIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQK 420

QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
Db 421 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 472

RESULT 6
US-10-281-319-4
; Sequence 4, Application US/10281319
; Publication No. US20030138820A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001002CON
; CURRENT APPLICATION NUMBER: US/10/281,319
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-397C
CURRENT APPLICATION NUMBER: US/10/188,186
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/303046
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/360814
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/303828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/323380
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/361133
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/304016
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/304502
PRIOR FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom
SEQ ID NO 76
LENGTH: 507
TYPE: PRT
ORGANISM: Homo sapiens
US-10-188-186-76

Query Match 87.1%; Score 2527.5; DB 12; Length 507;
Best Local Similarity 93.4%; Pred. No. 4.8e-227;
Matches 498; Conservative 0; Mismatches 2; Indels 33; Gaps 1;
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSGVPGRRRDKWSCSLVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSGVPGRRRDKWSCSLVAS 50
QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDITLLWSVTVSIFAIG 120
DB 51 -----YIKAFYNESWERRHRRPDPDITLLWSVTVSIFAIG 87
QY 121 GLVGTLLVQMTGKVLGRKHTLLANNFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
DB 88 GLVGTLLVQMTGKVLGRKHTLLANNFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 147
QY 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240
DB 148 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 207
QY 241 PAVVQLLSLPFLPDSPRYLLEKNEARAVKAFQTFGLKADVSQVEEVLAEASHVQSRIR 300
DB 208 PAVVQLLSLPFLPDSPRYLLEKNEARAVKAFQTFGLKADVSQVEEVLAEASHVQSRIR 267
QY 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFYNTSIFGKAGIPPAKIPYVTLST 360
DB 268 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFYNTSIFGKAGIPPAKIPYVTLST 327
QY 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPYLSIVGIL 420
DB 328 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPYLSIVGIL 387
QY 421 AIIASFCSGPGGIPFILTGEFFQOSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYC 480
DB 388 AIIASFCSGPGGIPFILTGEFFQOSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYC 447
QY 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
DB 448 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 500

APPLICANT: KEARNEY, Liam
APPLICANT: ELLIOTT, Vicky S.
APPLICANT: SEILHAMER, Jeffrey J.
APPLICANT: POLICKY, Jennifer L.
APPLICANT: BOROWSKY, Mark L.
APPLICANT: DING, Neil
APPLICANT: DING, Li
APPLICANT: LU, Dyung Aina M.
APPLICANT: HILLMAN, Jennifer L.
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
FILE REFERENCE: PI-0109 PCT
CURRENT APPLICATION NUMBER: US/10/297,022
CURRENT FILING DATE: 2002-11-25
PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747;
PRIOR FILING DATE: 2000-05-26; 2000-06-01; 2000-06-08; 2000-06-16; 2000-06-22; 2000-06-22
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PERL Program
SEQ ID NO 25
LENGTH: 537
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20030216310A1 7476643CD1
US-10-297-022-25

Query Match 93.8%; Score 2723; DB 15; Length 537;
Best Local Similarity 99.4%; Pred. No. 2.8e-245;
Matches 531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSGVPGRRRDKWSCSLVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSGVPGRRRDKWSCSLVAS 60
QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDITLLWSVTVSIFAIG 120
DB 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDITLLWSVTVSIFAIG 120
QY 121 GLVGTLLVQMTGKVLGRKHTLLANNFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
DB 121 GLVGTLLVQMTGKVLGRKHTLLANNFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240
DB 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240
QY 241 PAVVQLLSLPFLPDSPRYLLEKNEARAVKAFQTFGLKADVSQVEEVLAEASHVQSRIR 300
DB 241 PAVVQLLSLPFLPDSPRYLLEKNEARAVKAFQTFGLKADVSQVEEVLAEASHVQSRIR 300
QY 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFYNTSIFGKAGIPPAKIPYVTLST 360
DB 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFYNTSIFGKAGIPPAKIPYVTLST 360
QY 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPYLSIVGIL 420
DB 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPYLSIVGIL 420
QY 421 AIIASFCSGPGGIPFILTGEFFQOSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYC 480
DB 421 AIIASFCSGPGGIPFILTGEFFQOSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYC 480
QY 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 534
DB 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 534

RESULT 4
US-10-188-186-76
Sequence 76, Application US/10188186
Publication No. US20040029789A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.

; ORGANISM: Homo sapiens
US-10-276-774-2280

Query Match 94.3%; Score 2737; DB 12; Length 558;
Best Local Similarity 99.8%; Pred. No. 1.5e-246;
Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60
DB 19 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 78
QY 61 LAGAFGSSFLGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
DB 79 LAGAFGSSFLGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 138
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGGV 180
DB 139 GLVGTLLIVKMGKVLGRKHTLLANNPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGGV 198
QY 181 ALSVLPWYLSISPKIRSGLSQVTAIFICIGVFTGQLGLPELLGKESWYPLFGVIVV 240
DB 199 ALSVLPWYLSISPKIRSGLSQVTAIFICIGVFTGQLGLPELLGKESWYPLFGVIVV 258
QY 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300
DB 259 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 318
QY 301 LVSUVEILLRAPHYVRWQVTVVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
DB 319 LVSUVEILLRAPHYVRWQVTVVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 378
QY 361 GGIEETLAAVFSGLVIEHGRPLIIGGFLMGLPFGTITLTQLQDHAPWVPLSIVGIL 420
DB 379 GGIEETLAAVFSGLVIEHGRPLIIGGFLMGLPFGTITLTQLQDHAPWVPLSIVGIL 438
QY 421 AIIASFCSGPGGIPILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480
DB 439 AIIASFCSGPGGIPILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 498
QY 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
DB 499 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 551

RESULT 3

US-10-297-022-25
; Sequence 25, Application US/10297022
; Publication No. US20030216310A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: THORNTON, Michael
; APPLICANT: WALIA, Narinder K.
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Valda
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: RAUMANN, Brigitte E.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: GREENE, Barrie D.
; APPLICANT: KHAN, Farrah A.

LENGTH: 563 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 2;

US-09-981-947A-2

Query Match 100.0%; Score 2903; DB 9; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.7e-262;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60
QY 61 LAGAFGSSFLGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
DB 61 LAGAFGSSFLGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGGV 180
DB 121 GLVGTLLIVKMGKVLGRKHTLLANNPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGGV 180
QY 181 ALSVLPWYLSISPKIRSGLSQVTAIFICIGVFTGQLGLPELLGKESWYPLFGVIVV 240
DB 181 ALSVLPWYLSISPKIRSGLSQVTAIFICIGVFTGQLGLPELLGKESWYPLFGVIVV 240
QY 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300
DB 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300
QY 301 LVSUVEILLRAPHYVRWQVTVVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
DB 301 LVSUVEILLRAPHYVRWQVTVVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
QY 361 GGIEETLAAVFSGLVIEHGRPLIIGGFLMGLPFGTITLTQLQDHAPWVPLSIVGIL 420
DB 361 GGIEETLAAVFSGLVIEHGRPLIIGGFLMGLPFGTITLTQLQDHAPWVPLSIVGIL 420
QY 421 AIIASFCSGPGGIPILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480
DB 421 AIIASFCSGPGGIPILTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480
QY 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540
DB 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540
QY 541 TTPNTAWTQAAATTTATKKEHPL 563
DB 541 TTPNTAWTQAAATTTATKKEHPL 563

RESULT 2
US-10-276-774-2280
; Sequence 2280, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2280
; LENGTH: 558
; TYPE: PRT

Sequence 7, Appli
Sequence 65, Appl
Sequence 66, Appl
Sequence 7, Appli
Sequence 67, Appl
Sequence 64, Appl
Sequence 68, Appl
Sequence 10, Appl
Sequence 345, App
Sequence 15, Appl
Sequence 24, Appl
Sequence 2, Appli
Sequence 1248, Ap
Sequence 47, Appl
Sequence 48, Appl
Sequence 76, Appl
Sequence 3, Appli
Sequence 2, Appli
Sequence 50, Appl
Sequence 49, Appl
Sequence 6, Appli
Sequence 5, Appli
Sequence 51, Appl
Sequence 89, Appl
Sequence 124, App
Sequence 124, App
Sequence 125, App
Sequence 125, App
Sequence 122, App

16 1119.5 38.6 512 12 US-10-332-447-7
17 1112.5 38.3 502 12 US-10-029-020-65
18 1109.5 38.2 502 12 US-10-029-020-66
19 1108 38.2 500 9 US-09-981-947A-7
20 1103.5 38.0 502 12 US-10-029-020-67
21 1103.5 38.0 524 12 US-10-029-020-64
22 1099.5 37.9 501 12 US-10-029-020-68
23 1081 37.2 493 9 US-09-981-947A-10
24 1032.5 35.6 481 15 US-10-116-275-345
25 1000.5 34.5 499 12 US-10-332-447-15
26 997.5 34.4 483 12 US-10-029-020-24
27 979 33.7 524 13 US-10-094-059-2
28 936.5 32.3 492 15 US-10-295-027-1248
29 936 32.2 471 9 US-09-778-927A-47
30 935 32.2 501 9 US-09-778-927A-48
31 934.5 32.2 480 14 US-10-176-847-76
32 931.5 32.1 492 14 US-10-328-198-3
33 897.5 30.9 516 9 US-09-822-863-2
34 897.5 30.9 516 14 US-10-281-319-2
35 883 30.4 455 9 US-09-778-927A-50
36 879 30.3 471 9 US-09-778-927A-49
37 874 30.1 509 9 US-09-981-947A-6
38 866 29.8 494 9 US-09-981-947A-5
39 842 29.0 441 9 US-09-778-927A-51
40 830 28.6 496 15 US-10-341-434-89
41 773 26.6 522 15 US-10-099-323-124
42 773 26.6 522 15 US-10-044-564-124
43 772 26.6 522 15 US-10-099-322-125
44 772 26.6 522 15 US-10-044-564-125
45 767.5 26.4 523 15 US-10-099-322-122

ALIGNMENTS

RESULT 1
US-09-981-947A-2
Sequence 2, Application US/09981947A
Patent No. US20020164578A1
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/981,947A
FILING DATE: 18-Oct-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:18:33 ; Search time 46 Seconds
(without alignments)
3214.404 Million cell updates/sec

Title: US-09-981-947B-2
Perfect score: 2903
Sequence: 1 MARKQNRNSKGLVPLTDD.....NTAWTQAAATTATKKEHPL 563

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2903	100.0	563	9	US-09-981-947A-2
2	2737	94.3	558	12	US-10-276-774-2280
3	2723	93.8	537	15	US-10-297-022-25
4	2527.5	87.1	507	12	US-10-188-186-76
5	2396	82.5	472	9	US-09-822-863-4
6	2396	82.5	472	14	US-10-281-319-4
7	2341	80.6	480	12	US-10-188-186-80
8	2337.5	80.5	565	12	US-10-188-186-78
9	2017	69.5	422	12	US-10-188-186-90
10	1939	66.8	416	14	US-10-168-651-6
11	1929	66.4	396	12	US-10-188-186-86
12	1920	66.1	396	12	US-10-188-186-84
13	1911	65.8	396	12	US-10-188-186-88
14	1721	59.3	362	12	US-10-188-186-82
15	1122.5	38.7	501	15	US-10-295-027-308

Db	10	LIPAI	SIATIGSFQFYNTGVINAPEAIKOPFNLTLEERSETPSPSSVLLTSLMSLSVAI	69
QY	117	FAIGGLVGLIYKMGKVLGRKHTLLANN	FAISAAALMACSLQAGAPEMLIVGRFTMGI	176
Db	70	FVGGMIGSFVGLFVNFGRNSMLIVNL	AIAGGCLMGFCKIAESVEMILGRLLIGL	129
QY	177	DGGVALSVLPMYLSISPKIRGSLGQVTA	IFICIGVFTGQLLGPHELLGKESWPLYFG	236
Db	130	FCGLCTGFVPWYIGISPTALRGAFTLN	QLGIVIGILVAQIFGLKVLGTEDLWPLLLG	189
QY	237	VIVPAAVOLLSLPFLPDSPRYLLEKNE	ARAVAFQTFGLKADVSQOEVEEVLAEHVQ	296
Db	190	FTILPAIIQCAALPFCPSPRELLINRKE	EKAKEILQRLWGTEDVAQIOEMKDESNRM	249
QY	297	RSIRLVSVLELARAPVVRQVVTVITW	MACYQLOGLNAIWEFTWSIFCKAGIPPAKIPYV	356
Db	250	SOEKQVTVLELFRAPNYQPIIISIM	LOLSQOLSGINAVFYISTGIFKADGV--QEPVYA	307
QY	357	TLSTGGIETLAAVFSGLVIEHLGRPL	LIGGFGLMGFPFGTLITTLIQDHAPWVPYLSI	416
Db	308	TIGAGVWNTIFTVSVFLVERAGRTLH	LGLGGWAFCSILMTISLLKONYSWMSFICI	367
QY	417	VGILAIIASFCGPGGIPFILTGEFFQ	SQSRPAAFIAGTVNMLSNFAVGLLPFFIOKSL	476
Db	368	GAILVFAFFETIGPGPIFWFIVAELF	QGQPRPAAAMAVAGCSNWTSNFLVGLLPFSATFYL	427
QY	477	DTYCFVETATICTGAIYLYFVLPET	KNTYAEISQAF	514
Db	428	GAYFIVFTVFLVIFWVFTFFKVPET	GRTEETRAF	465

Search completed: April 6, 2004, 10:19:34
Job time : 24 secs

QY 293 SHVQSRIRLVLELLRAPYVRWQVTVITMACYQLCGLNAINFYTNISIFGKAGI-PPA 351
Db 264 KRKLREPLSLQLLQSGRTHRQPLIIAVVLQSLQSGINAVFYYSIFETAGVGQPA 323
QY 352 KIPYVTLSTGGIETLAASFGLVIEHLGRPL-LIGGFLMGLFGTTLITLTLQDHAPW 410
Db 324 ---YATIGAGVNTVFTLVSVLLVERAGRTLHLGLAGMCGCAI-LMTVALLLERVPA 379
QY 411 VPYLSIVGLIATIASPCSGPGIPILTCGRFPQOSORPAAFIAGTVNVLNENAVGLLFP 470
Db 380 MSYVSIVAIFGVAFPEIGPGIPFWIVAEFLFSQGRPAAMAVAGFSNWTNFIIGMGFQ 439
QY 471 FIKSLDTCFLVFATICTITGAIYLVFLVLPETKNRYAEISQAFSKRNKAYPPEEK 526
Db 440 YVAEAMGPYVFLFAVLLGFFITFLRVPEIRGRTFDQISAFHRTPSLLEQEVK 495

RESULT 14

US-09-610-417-6
; Sequence 6, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-610-417-6

Query Match 30.1%; Score 874; DB 4; Length 509;
Best Local Similarity 38.7%; Pred. No. 3.4e-78;
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;

QY 57 LVASLAGAGSSFLYGNLSVNNAPTPIKAFYNESWERRHG----RPIDPDTLLMSV 112
Db 24 LVLAFAVLSLQFGYGNIGVINAPQKVIQSQNETWLGRQPGEGSSIPGPTLTLAL 83
QY 113 TVSIFAIGLVGTLIVMTGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLVGRF 172
Db 84 SVAIFSVGGMISSFLIGIISQWLGRKRAMLVNVLAVGLGSLMGLANAAASYEMILIGRF 143

QY 173 IMGIDGVALSVLPMYLSISPKIRGSLGQVTAIFCIGVFTGQLLGLPELLGKESTWP 232
Db 144 LIGAYSLTSGLVPMYVGTIAPHLRGALGTNLQLAIVIGILIAQVLGLESLLGTASLWP 203
QY 233 YLFGVIVPAVWQLLSIPLPDPSRVLLEKNEARAVKAFOTFLGKADVSQREBVLA 292
Db 204 LLGLTVPALLOLVLLPQCPSPRYLYTIQNLGEPARKSLKRLTGWADVSGVLAELKDE 263
QY 293 SHVQSRIRLVLELLRAPYVRWQVTVITMACYQLCGLNAINFYTNISIFGKAGI-PPA 351
Db 264 KRKLREPLSLQLLQSGRTHRQPLIIAVVLQSLQSGINAVFYYSIFETAGVGQPA 323
QY 352 KIPYVTLSTGGIETLAASFGLVIEHLGRPL-LIGGFLMGLFGTTLITLTLQDHAPW 410
Db 324 ---YATIGAGVNTVFTLVSVLLVERAGRTLHLGLAGMCGCAI-LMTVALLLERVPA 379
QY 411 VPYLSIVGLIATIASPCSGPGIPILTCGRFPQOSORPAAFIAGTVNVLNENAVGLLFP 470
Db 380 MSYVSIVAIFGVAFPEIGPGIPFWIVAEFLFSQGRPAAMAVAGFSNWTNFIIGMGFQ 439
QY 471 FIKSLDTCFLVFATICTITGAIYLVFLVLPETKNRYAEISQAFSKRNKAYPPEEK 526
Db 440 YVAEAMGPYVFLFAVLLGFFITFLRVPEIRGRTFDQISAFHRTPSLLEQEVK 495

RESULT 15

US-09-031-392-5
; Sequence 5, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-5

Query Match 29.8%; Score 866; DB 2; Length 494;
Best Local Similarity 38.9%; Pred. No. 2e-77;
Matches 178; Conservative 90; Mismatches 188; Indels 2; Gaps 1;

QY 57 LVASLAGAGSSFLYGNLSVNNAPTPIKAFYNESWERRHGRIIDPDTLLMSVTSI 116
Db 57 LVASLAGAGSSFLYGNLSVNNAPTPIKAFYNESWERRHGRIIDPDTLLMSVTSI 116

Db 429 CGPVFIIFVLLVLFVFIIRTYKVPETKGRTPDEIASGFRQGGASQDKT--PSE 481

RESULT 12

US-09-031-392-6

Sequence 6, Application US/09031392

Patent No. 5942398

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Weng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES

TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,392

FILING DATE: 26-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 509 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-031-392-6

Query Match 30.1%; Score 874; DB 2; Length 509;

Best Local Similarity 38.7%; Pred. No. 3.4e-78;

Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;

Qy 57 LVASLAGAGSSFLYGNLSVNVNAPPYIKAFYNESWERRHG-----RPIDDTITLLMSV 112

Db 24 LVAVFSVAVLSLQFGYGNIGVINAPQKVIQSYNETWLGROGPEGPSSIPPGTTLTLLWAL 83

Qy 113 TVSIFAIGGLVGTLLVIMKIGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGRF 172

Db 84 SVAIFSVMGMISSFLIGISQWLGKRAMLVNNVAVLGGSLMGSLGANAASYLEMLILGRF 143

Qy 173 IMGIDGGVALSVLPMYLSLSEIPKIRGSLGVTAIFICIGVFTGQLLGLPGLLKESTWP 232

Db 144 LIGASGLTSLGLVPMYVGEIAPHLAGALGTNLQALVIGILIAQVGLSLLGTASLWP 203

Qy 233 YLFGVIVPVAVQLLSPLDPSPRVYLLEKNEARAVKAFOTFGLKADVQSEVEVLAE 292

Db 204 LLGLGLTLPALLQVLLPFCPSPRYLIYIQNLEGPARKSLKRLTGMADVSGVLAELKDE 263

Qy 293 SHVORSTRVLSVLELLRAPVWQVVTIVTWACVQLCLNAIMFTVNSIFGKAGI-PPA 351

Db 264 KRKLREPRLSLQLLGSRTHQPLIIAVVQLSQQSLGINAVFYYSIFETAGVGQPA 323

Qy 352 KIPYVTLSTGGIETLAASVGLVIEHLGRPL-LIGGFLGMLGFTGTITITLTQDRAPV 410

Db 324 ---YATIGAGVNVNTVTLVSVLLVERGRTTLHLGLAGMCCAI-LMTVALLLLIERVPA 379

Qy 411 VPYLSIVGILLAIASFCSPGGIPILTGTBFQSQRPAAPIAGTVNVLNFAVGLLFP 470

Db 380 MSYVSIVAIFGVFAVFEIGPGPIFWFVIAELFSQPRPAAVAVAGFSNWTSTNFIIGMPQ 439

Qy 471 FIQKSLDTCFLAVENATICITGAIVLYFVLPEKRTVAELISQAFSKENKAYPPEEK 526

Db 440 YVAAWGPVYVLLFAVLLGFFITFLRVPSTGRFTDQISAAFHRTPSLLEQEVK 495

RESULT 13

US-09-299-549-6

Sequence 6, Application US/09299549

Patent No. 6136547

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Weng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

TITLE OF INVENTION: GLUTEX AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,549

FILING DATE: 26-APR-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/031,392

FILING DATE: 26-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 509 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-299-549-6

Query Match 30.1%; Score 874; DB 3; Length 509;

Best Local Similarity 38.7%; Pred. No. 3.4e-78;

Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;

Qy 57 LVASLAGAGSSFLYGNLSVNVNAPPYIKAFYNESWERRHG-----RPIDDTITLLMSV 112

Db 24 LVAVFSVAVLSLQFGYGNIGVINAPQKVIQSYNETWLGROGPEGPSSIPPGTTLTLLWAL 83

Qy 113 TVSIFAIGGLVGTLLVIMKIGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGRF 172

Db 84 SVAIFSVMGMISSFLIGISQWLGKRAMLVNNVAVLGGSLMGSLGANAASYLEMLILGRF 143

Qy 173 IMGIDGGVALSVLPMYLSLSEIPKIRGSLGVTAIFICIGVFTGQLLGLPGLLKESTWP 232

Db 144 LIGASGLTSLGLVPMYVGEIAPHLAGALGTNLQALVIGILIAQVGLSLLGTASLWP 203

Qy 233 YLFGVIVPVAVQLLSPLDPSPRVYLLEKNEARAVKAFOTFGLKADVQSEVEVLAE 292

Db 204 LLGLGLTLPALLQVLLPFCPSPRYLIYIQNLEGPARKSLKRLTGMADVSGVLAELKDE 263

NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE: Human
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..492
OTHER INFORMATION: Facilitative glucose transporter
OTHER INFORMATION: Glut1 protein
PCT-US95-16126-3

STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/355,844
FILING DATE: 14-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE: Human
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..492
OTHER INFORMATION: Facilitative glucose transporter
OTHER INFORMATION: Glut1 protein
PCT-US95-16126-3

Query Match 31.7%; Score 919.5; DB 2; Length 492;
Best Local Similarity 39.2%; Pred. No. 9.4e-83;
Matches 186; Conservative 98; Mismatches 180; Indels 11; Gaps 5;

QY 57 LVASLAGAGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPDPTLTLLWSVTYSI 116
DB 12 LMLAVGAVLGLSQFGYNTGVINAPQKVBEEFYNTWVHRYGESILPTLTLLWSLVAI 71
QY 117 FAIGGLVGLTLIVKMGKVLGRKHTLLANNGFSAIALLMACSLQAGAFEMLIIVGRFIMG 176
DB 72 FSVGGMIGSFVGLFVNRFGRRNSMLMNLALLAFVSALVLMGFSKLGKSFEMILGRFIIV 131
QY 177 DGGVALSVLPMYLSIEISPKIRSGLVQVTAIFICIGVFTGQLLGPPELLGKSTWPLYFG 236
DB 132 YCGLTTGFPVMYGEVSPTAFRGALGTLHQLGIIVVGILIAQVGLDSIMGNDLWPLLS 191
QY 237 VIIVPAVQVLLSLPFLDPSRYPYLLEKNEARAVKAFQTLGKADVSQVEEVLAEHVQ 296
DB 192 IIFIPALLQCVIPFCPSPRFLINRNEENRAKSVLKKLGTADVTDLQEMKEESRQM 251
QY 297 RSIRLVSVLELLRAPYVRQVTVITMACYQLCGLNATWFTNSIFGKAGIPPAKIPYV 356
DB 252 MREKVTILELFRSPAYRQPIILIAVVLQLSQQUSGINAVFYSTSIPEKAGV--QQPYVA 309
QY 357 TLSTGGIETLAAVFSGLVIEHLGRRL-LIGGFLMGLFFGTITLTITLQDHAPVVPYLS 415
DB 310 TIGSGIVNTAFTVSLFVVERAGRTLHLIGLAGMAGQAI-LMTIALALLEQLPWSYLS 368
QY 416 IVGILAIASFCSGPGGIPFLTGEFFQSQRPAAFIAGTVNWSNFVGLLFPFIQKS 475
DB 369 IVAIFGVAFVFEVGPPIWFIVAELESQGPRAIAVAGFSNWTNSNFIWCMCFQVVEQL 428
QY 476 LDTYCFVLFATICTGAIYLYFVLPETKNTVABISQAF-----SKNKAYPPEE 525
DB 429 CGPVYFIITVLLVLFIRTYFKVPETKRTDEIASGFRQGSQSDXT--PEE 481

RESULT 11
PCT-US95-16126-3
Sequence 3, Application PC/TUS9516126
GENERAL INFORMATION:
APPLICANT: Fischberg, Jorge
APPLICANT: Czegledy, Ferenc
APPLICANT: Iserovich, Pavel
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
STRUCTURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

Query Match 31.7%; Score 919.5; DB 2; Length 492;
Best Local Similarity 39.2%; Pred. No. 9.4e-83;
Matches 186; Conservative 98; Mismatches 180; Indels 11; Gaps 5;

QY 57 LVASLAGAGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPDPTLTLLWSVTYSI 116
DB 12 LMLAVGAVLGLSQFGYNTGVINAPQKVBEEFYNTWVHRYGESILPTLTLLWSLVAI 71
QY 117 FAIGGLVGLTLIVKMGKVLGRKHTLLANNGFSAIALLMACSLQAGAFEMLIIVGRFIMG 176
DB 72 FSVGGMIGSFVGLFVNRFGRRNSMLMNLALLAFVSALVLMGFSKLGKSFEMILGRFIIV 131
QY 177 DGGVALSVLPMYLSIEISPKIRSGLVQVTAIFICIGVFTGQLLGPPELLGKSTWPLYFG 236
DB 132 YCGLTTGFPVMYGEVSPTAFRGALGTLHQLGIIVVGILIAQVGLDSIMGNDLWPLLS 191
QY 237 VIIVPAVQVLLSLPFLDPSRYPYLLEKNEARAVKAFQTLGKADVSQVEEVLAEHVQ 296
DB 192 IIFIPALLQCVIPFCPSPRFLINRNEENRAKSVLKKLGTADVTDLQEMKEESRQM 251
QY 297 RSIRLVSVLELLRAPYVRQVTVITMACYQLCGLNATWFTNSIFGKAGIPPAKIPYV 356
DB 252 MREKVTILELFRSPAYRQPIILIAVVLQLSQQUSGINAVFYSTSIPEKAGV--QQPYVA 309
QY 357 TLSTGGIETLAAVFSGLVIEHLGRRL-LIGGFLMGLFFGTITLTITLQDHAPVVPYLS 415
DB 310 TIGSGIVNTAFTVSLFVVERAGRTLHLIGLAGMAGQAI-LMTIALALLEQLPWSYLS 368
QY 416 IVGILAIASFCSGPGGIPFLTGEFFQSQRPAAFIAGTVNWSNFVGLLFPFIQKS 475
DB 369 IVAIFGVAFVFEVGPPIWFIVAELESQGPRAIAVAGFSNWTNSNFIWCMCFQVVEQL 428
QY 476 LDTYCFVLFATICTGAIYLYFVLPETKNTVABISQAF-----SKNKAYPPEE 525
DB 429 CGPVYFIITVLLVLFIRTYFKVPETKRTDEIASGFRQGSQSDXT--PEE 481

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-610-417-10

Query Match 37.2%; Score 1081; DB 4; Length 493;
Best Local Similarity 48.5%; Pred. No. 7.8e-99;
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;

QY 55 SLLVASLAGAGSSFLYGYNLVVNAPTPIKAFYNESWERRHGRPIDDTITLLWSVTV 114
DB 13 TLVLAVLIAALG-SFOYGYNLGVINAPQKVEAFY-ETWLGXGEXPSVPTITLLWSLSV 70
QY 115 SIFAIGGLVGLTIIVKMIGKVLGRKHTLLANNPFAISALLMACSLQAGAFEMLIIVGRFIM 174
DB 71 SIFAVGGMIGSFLVXIGNRLGRKXAMLVNVLAIAGGLMGLAKAXSFEMLIIGRFII 130
QY 175 GIDGGVALSVLPMYLSPEISPEIRSGISGOVTAIFICIGVFTGQLLPELLGKESTWPLY 234
DB 131 GLYCLSGGVPMYVGEISPTALRGALTNQLGIVIGILIAQVLGDSLLGNESLWPLL 190
QY 235 FGVIVPAAVQVLLSLPPLDPSRYLLLEKHNKNEARAVKAFQTLGKADVQSEVEVLAESH 294
DB 191 LGTGVPAQLQLLLPFCPSFRYLLINKNEARAKKALQRLGTADVQSEVAEMKDESR 250
QY 295 VORSIRLVSVLELRAPYVRQWVTVVITMACYQLCGLNAINWFTYNSIFGKAGIPPAKIP 354
DB 251 XMKSEKXSVLELFRSRYRQPVIIAIVLQSLQSGINAVFYSTISFEKAGV--GQPV 308
QY 355 YVTLSTGIGTETLAIVFSLGIVIEHLGRRELLIGGGLMGLFEGTLTITLQDHAPWVPLY 414
DB 309 YATIGAGVNVTVFTVSVFVERAGRRTLLHLLGGMAGCAVMTIALALDQVPMWSYV 368
QY 415 STVGILAIASFCSGPGGIPILTGEFFQOSQORPAAFIIAGTVNWLNSFAVGLLPPFIQK 474
DB 369 SIVAIFGVFAFFVGGPIPMFIVAEFLSQGPRPAATAVAGFSNWTNSNFIIVGLLFQYIAE 428
QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
DB 429 LLGPYFVIVFAVLLLLFFITFLKVPETKGRTFDEIAAFKRNKXEQPEKE 480

RESULT 10
US-08-355-844-3
Sequence 3, Application US/08355844
Patent No. 5940307
GENERAL INFORMATION:
APPLICANT: Fischbarg, Jorge
APPLICANT: Czegledy, Ferenc
APPLICANT: Iserovich, Pavel
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
TITLE OF INVENTION: STRUCTURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donchue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/355,844
FILING DATE: 14-DEC-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:

QY 55 SLLVASLAGAGSSFLYGYNLVVNAPTPIKAFYNESWERRHGRPIDDTITLLWSVTV 114
DB 13 TLVLAVLIAALG-SFOYGYNLGVINAPQKVEAFY-ETWLGXGEXPSVPTITLLWSLSV 70
QY 115 SIFAIGGLVGLTIIVKMIGKVLGRKHTLLANNPFAISALLMACSLQAGAFEMLIIVGRFIM 174
DB 71 SIFAVGGMIGSFLVXIGNRLGRKXAMLVNVLAIAGGLMGLAKAXSFEMLIIGRFII 130
QY 175 GIDGGVALSVLPMYLSPEISPEIRSGISGOVTAIFICIGVFTGQLLPELLGKESTWPLY 234
DB 131 GLYCLSGGVPMYVGEISPTALRGALTNQLGIVIGILIAQVLGDSLLGNESLWPLL 190
QY 235 FGVIVPAAVQVLLSLPPLDPSRYLLLEKHNKNEARAVKAFQTLGKADVQSEVEVLAESH 294
DB 191 LGTGVPAQLQLLLPFCPSFRYLLINKNEARAKKALQRLGTADVQSEVAEMKDESR 250
QY 295 VORSIRLVSVLELRAPYVRQWVTVVITMACYQLCGLNAINWFTYNSIFGKAGIPPAKIP 354
DB 251 XMKSEKXSVLELFRSRYRQPVIIAIVLQSLQSGINAVFYSTISFEKAGV--GQPV 308
QY 355 YVTLSTGIGTETLAIVFSLGIVIEHLGRRELLIGGGLMGLFEGTLTITLQDHAPWVPLY 414
DB 309 YATIGAGVNVTVFTVSVFVERAGRRTLLHLLGGMAGCAVMTIALALDQVPMWSYV 368
QY 415 STVGILAIASFCSGPGGIPILTGEFFQOSQORPAAFIIAGTVNWLNSFAVGLLPPFIQK 474
DB 369 SIVAIFGVFAFFVGGPIPMFIVAEFLSQGPRPAATAVAGFSNWTNSNFIIVGLLFQYIAE 428
QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
DB 429 LLGPYFVIVFAVLLLLFFITFLKVPETKGRTFDEIAAFKRNKXEQPEKE 480

RESULT 9
US-09-610-417-10
Sequence 10, Application US/09610417
Patent No. 6346374
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
NUCLEOTIDE SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/610,417
FILING DATE: 05-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/299,549
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids

QY 236 GTIVFAVQLLSLPDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 295
DB 197 GTGVPAAQLQLLPFPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 256
QY 296 QRSIRVSVLELLRABPYVQVVTIVTMACYQLCGNLAIWFTYNSIFGKAGIPPAKIPY 355
DB 257 ERAAGFISVLKFRMSLAWQLLSIIIVLGGQQLSGVNAIYYAQIYLSAGVPEEHVQY 316
QY 356 VTLSTGGIE---TLAIVFSLVIEHLGRPLIGGFLMGLPFGTLTITLTLODHAPWVP 412
DB 317 VTAGTGANVWMTFCFAVF---VVELLGRLLLLGFSICLIACCVLTAALALQDTVSWMP 373
QY 413 YLSIVGILAIASFCSPGGIPFLTGEFFQOSQORPAAFIIAGTVNWLNSFVAVGLLPFI 472
DB 374 YLSIVGILAIASFCSPGGIPFLTGEFFQOSQORPAAFIIAGTVNWLNSFVAVGLLPFI 432
QY 473 QKSLDYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNK---AYPPEKI 527
DB 433 QKSLDYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNK---AYPPEKI 490

RESULT 7

US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-10

Query Match 37.2%; Score 1081; DB 2; Length 493;
Best Local Similarity 48.5%; Pred. No. 7.8e-99;
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;
QY 55 SILVSLAAGAFSSFLYGNLSVWNAPTYIKAFYNESWERHGRPIDPDTLLTILMSVTV 114
DB 13 TLVLAVLIAALG-SFOYGNLGVINAPQKVIEAFY-ETWLGRGXEPSPTLLTILMSLSV 70
QY 115 SIFAIGLGVTLVKMIGKVLGRKHTLLANNGFATSAALLMACSLQAGAFEMLVGRFTM 174
DB 71 SIFAVGMLGSLVIGNIGRLGRKXAMVNVNVAIAGGLMGLAXKXAFEMLVGRFTM 130

QY 175 GIDGGVALSVLPMYLSIPKSEIRGSLGQVTAIFICIGVFTGQLLGPPELLGKESTWPLY 234
DB 131 GLYCLSGGVPMVGEISPTALRGALGTNLQIGIVIGILIAQVLGDLSDLLGNESLWPLL 190
QY 235 FGVIWFAVQQLSLPDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 294
DB 191 LGLTGVPALLQLLPFPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 250
QY 295 VQRIRVSVLELLRABPYVQVVTIVTMACYQLCGNLAIWFTYNSIFGKAGIPPAKIP 354
DB 251 XMXSEKVSULELFRSRKXQPVIIAIVLQSLQSGINAVFYVYSTSIFPKAGV--QGPV 308
QY 355 YVTLSTGGIEFLAIVFSLVIEHLGRPLIGGFLMGLPFGTLTITLTLODHAPWVP 414
DB 309 YATIGAGVNVTVFVSVFVVERAGRRLTLHLGLGGMAGCAVLMTIALALLDQVPMWSYV 368
QY 415 SIVGILAIASFCSPGGIPFLTGEFFQOSQORPAAFIIAGTVNWLNSFVAVGLLPFIQK 474
DB 369 SIVAIFGFAVFEVGPDPFIPWIVAEIIFSQGPRAATAVAGFSNWTNFIIVGLLFOYIAE 428
QY 475 SLDTYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNKAYPPEEK 526
DB 429 LLGPVVFIVFAVLLLLFFITFLKVPETKGRTPDEIAAFAFKKXKXQPEKE 480

RESULT 8

US-09-299-549-10
; Sequence 10, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-10

Query Match 37.2%; Score 1081; DB 3; Length 493;
Best Local Similarity 48.5%; Pred. No. 7.8e-99;
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;

Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMKNVSEVYPEKEEL 490

RESULT 5

US-09-299-549-7

Sequence 7, Application US/09299549

Patent No. 6136547

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Weng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

TITLE OF INVENTION: GLUTEX AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FASTSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,549

FILING DATE: 26-APR-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/031,392

FILING DATE: 26-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 500 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-299-549-7

Query Match 38.2%; Score 1108; DB 3; Length 500;

Best Local Similarity 46.2%; Pred. No. 1.6e-101;

Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;

Qy 56 LLVASLAGAGSSFLYGYNLSVNAAPTPIKAFYNESWERRHGRPIDPDTLLTLLMSVTVS 115

Db 17 LALATLIAAFSGSFQYGVNVAVNSPALLMQQFYNETYYGRTGFEFMDPFLTLLMSVTVS 76

Qy 116 IFATGLVGTLLVKNMGKVLGRKHTLLANNGFALSAALLMACSLQAGAFEMLIIVGRFIMG 175

Db 77 MFPPGFGTIGSLVGLVPLNKNFGKRGALLFNFIIVPALMGCSRVATSFELIIISRLVVG 136

Qy 176 IDGGVALSVLPMYLSSEIPKEIRSGVTAIFTCIGVFTGQLLGLPELLGKESWPYLF 235

Db 137 ICAGVSNVPMYLGELAPKNLRGALGVWPQLFTVIGILVAQIFGLNLLANVDGWPILL 196

Qy 236 GVIIVPAVQLSLPFLPDSPPRYLLLEKNEARAVKAFQTFGLKADVSQREVEVLAESHV 295

Db 197 GLTGVPAAQLQLLPPFPESPRLYLIQKDBAAAKALQTLRGWDSVDREVAERQDEA 256

Qy 296 QRSRLVSVLELLKAPYVRQVTVIVTMACYQLCGLNATWFTNSIFGAGIPPAKIPY 355

Db 257 EKAAGFISVLKFRMSRLRWQLLSIIILMGQQQLSGVNAIYYADQIYLSAGVPEEHVQY 316

Qy 356 VTLSTGGIE---TLAAVFSGLVIEHLGRRPLLIGGFGMLGFFGCTLTITLTLDQADPWP 412

Db 317 VTAGTGVNVMTFCAVF---VVELLGRLLLLGFGICLIACCVLTAALALQDTVSWMP 373

Qy 413 YLSIVGILAIATFSCSGPGGIPFILTGBEFFQOORPAAFIAGTVNLSNFAVGLLPFFI 472

Db 374 YISIVCVISYVIGHALGPSIPALLI-TIFLQSRPSAFWVGSVHLSNFTVGLLPFFI 432

Qy 473 QKSLDTYCYFLVFATICTITGAIYLYFLVLPETKNTYAEISQAFSKRNK---AYPEEKI 527

Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMKNVSEVYPEKEEL 490

RESULT 6

US-09-610-417-7

Sequence 7, Application US/09610417

Patent No. 6346374

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Weng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

TITLE OF INVENTION: GLUTEX AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FASTSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/610,417

FILING DATE: 05-JUL-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/299,549

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 500 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-610-417-7

Query Match 38.2%; Score 1108; DB 4; Length 500;

Best Local Similarity 46.2%; Pred. No. 1.6e-101;

Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;

Qy 56 LLVASLAGAGSSFLYGYNLSVNAAPTPIKAFYNESWERRHGRPIDPDTLLTLLMSVTVS 115

Db 17 LALATLIAAFSGSFQYGVNVAVNSPALLMQQFYNETYYGRTGFEFMDPFLTLLMSVTVS 76

Qy 116 IFATGLVGTLLVKNMGKVLGRKHTLLANNGFALSAALLMACSLQAGAFEMLIIVGRFIMG 175

Db 77 MFPPGFGTIGSLVGLVPLNKNFGKRGALLFNFIIVPALMGCSRVATSFELIIISRLVVG 136

Qy 176 IDGGVALSVLPMYLSSEIPKEIRSGVTAIFTCIGVFTGQLLGLPELLGKESWPYLF 235

Db 137 ICAGVSNVPMYLGELAPKNLRGALGVWPQLFTVIGILVAQIFGLNLLANVDGWPILL 196

US-09-031-392-7
 ; Sequence 7, Application US/09031392
 ; Patent No. 5942398
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Weng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 500 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-031-392-7

Query Match 38.2%; Score 1108; DB 2; Length 500;
 Best Local Similarity 46.2%; Pred. No. 1.6e-101;
 Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;

QY	56	LLVASLAGAFSSFLYGNLSVNVNAPTYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVS	115
Db	17	LALATLIAAFGSSFOYGYNVAANSPALLMQOFTYGTGTGTFDFPDTLLTLLWSVTVS	76
QY	116	IFATGGLVGLTVKMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFMG	175
Db	77	MFPFGFTGSLVGLVPLVKNFKRGKALLFNNTFSIVPAILMCCSRVATSFELIIISRLIVG	136
QY	176	IDGGVALSVLPMYLSLSEISPKIRSGISLGOVTAIFICIGVFTGQLLGLPELLGKSTWPLYF	235
Db	137	ICAGVSSNVVPMYLGELAPKNLRGALGVVPOLFITVGLVLAQIFGLRNLLANVDGWPILL	196
QY	236	GVIVVPAVVQLLSLFPDPSPRYLLLEKHNARAVKAFQTLGKADVSQVEEVLAESHV	295
Db	197	GLTGVPALQQLLLFPFPESPRYLLIQKDEAAKALQTLRGWDSVDREVAIRQDEA	256
QY	296	QRSLRLVSVLELLRAPYRWQVTVITMACYQGLNAINWFYTNISFGKAGIPAKIPY	355
Db	257	EKAAGFISVGLKLFMRSLRWQLLSIIVLMGGQGLSGVNAIYYADQIYLSAGVPEEHVQY	316
QY	356	VTLSGGE---TLAAVPSGLVIEHGLRPLLIGFGLMGLFFGTTLTTLQDHAPWVP	412
Db	317	VTAGTGVNVVMTFCAVF---VVELLGRLLLLGLGFSICLIACCVLTAALALQDTSWMP	373
QY	413	YLSIVGILAIASFCSPGGIPFILTGFQQSORPAAFIAGTVNMLSNFVAGLLPFPFI	472
Db	374	YISIVCVISVIGHALGSPDALLI-TIFQSSRPSAFMVGSGVHMLSNVTGLIFPFI	432
QY	473	QKSLDTCFLVPATICITGAIYLYFVLPETKNTVAEISQAFSKRNK---AYPPEKI	527

COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/610,417
 FILING DATE: 05-JUL-2000
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/299,549
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Meiklejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 563 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-610-417-2

Query Match 100.0%; Score 2903; DB 4; Length 563;
 Best Local Similarity 100.0%; Pred. No. 3.9e-280;
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRKQWNSCSILLVAS	60
Db	1	MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRKQWNSCSILLVAS	60
QY	61	LACAFSSFLYGNLSVNVNAPTYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSFAIG	120
Db	61	LACAFSSFLYGNLSVNVNAPTYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSFAIG	120
QY	121	GLVGTLLVVMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFMGIDGV	180
Db	121	GLVGTLLVVMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFMGIDGV	180
QY	181	ALSVLPMYLSLSEISPKIRSGISLGOVTAIFICIGVFTGQLLGLPELLGKSTWPLYFVIV	240
Db	181	ALSVLPMYLSLSEISPKIRSGISLGOVTAIFICIGVFTGQLLGLPELLGKSTWPLYFVIV	240
QY	241	PAVVQLLSLFPDPSPRYLLLEKHNARAVKAFQTLGKADVSQVEEVLAESHVORSTR	300
Db	241	PAVVQLLSLFPDPSPRYLLLEKHNARAVKAFQTLGKADVSQVEEVLAESHVORSTR	300
QY	301	LVSVLELLRAPYRWQVTVITMACYQGLNAINWFYTNISFGKAGIPAKIPYVTLST	360
Db	301	LVSVLELLRAPYRWQVTVITMACYQGLNAINWFYTNISFGKAGIPAKIPYVTLST	360
QY	361	GGIETLAAVPSGLVIEHGLRPLLIGFGLMGLFFGTTLTTLQDHAPWVPYLSIVGIL	420
Db	361	GGIETLAAVPSGLVIEHGLRPLLIGFGLMGLFFGTTLTTLQDHAPWVPYLSIVGIL	420
QY	421	AIIASFCSGGGIPFILTGFQQSORPAAFIAGTVNMLSNFVAGLLPFPFIQKSLDTC	480
Db	421	AIIASFCSGGGIPFILTGFQQSORPAAFIAGTVNMLSNFVAGLLPFPFIQKSLDTC	480
QY	481	FLVPATICITGAIYLYFVLPETKNTVAEISQAFSKRNKAYPPEEKIDSAVTPASSPF	540
Db	481	FLVPATICITGAIYLYFVLPETKNTVAEISQAFSKRNKAYPPEEKIDSAVTPASSPF	540
QY	541	TTNTNAMIQAATTTATKKEHPL	563
Db	541	TTNTNAMIQAATTTATKKEHPL	563

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 563 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-299-549-2

Query Match      100.0%; Score 2903; DB 3; Length 563;
Best Local Similarity 100.0%; Pred. No. 3.9e-280;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLTLLWSVTVSIFAIG 120
Db 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
Db 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFGVIIV 240
Db 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFGVIIV 240
QY 241 PAVQQLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEASHVORSIR 300
Db 241 PAVQQLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEASHVORSIR 300
QY 301 LVSVLELLRAPYVRWQVTVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
Db 301 LVSVLELLRAPYVRWQVTVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
QY 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLTTLTLODHAPWVPYLSIVGIL 420
Db 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLTTLTLODHAPWVPYLSIVGIL 420
QY 421 AIIASFCSGPGGIPFILTGBFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480
Db 421 AIIASFCSGPGGIPFILTGBFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480
QY 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540
Db 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540
QY 541 TTPNTAWIQAAATTTATKKEHPL 563
Db 541 TTPNTAWIQAAATTTATKKEHPL 563

```

RESULT 2

```

US-09-299-549-2
; Sequence 2, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 563 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-299-549-2

Query Match      100.0%; Score 2903; DB 3; Length 563;
Best Local Similarity 100.0%; Pred. No. 3.9e-280;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLLECDHLRSQVPGRRKDWSCSLIVAS 60
Db 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLLECDHLRSQVPGRRKDWSCSLIVAS 60
QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLTLLWSVTVSIFAIG 120
Db 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
Db 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFGVIIV 240
Db 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFGVIIV 240
QY 241 PAVQQLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEASHVORSIR 300
Db 241 PAVQQLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEASHVORSIR 300
QY 301 LVSVLELLRAPYVRWQVTVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
Db 301 LVSVLELLRAPYVRWQVTVITMACYQLCGLNAINFWYTNISIFGKAGIPPAKIPYVTLST 360
QY 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLTTLTLODHAPWVPYLSIVGIL 420
Db 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLTTLTLODHAPWVPYLSIVGIL 420
QY 421 AIIASFCSGPGGIPFILTGBFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480
Db 421 AIIASFCSGPGGIPFILTGBFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480
QY 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540
Db 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540
QY 541 TTPNTAWIQAAATTTATKKEHPL 563
Db 541 TTPNTAWIQAAATTTATKKEHPL 563

```

RESULT 3

```

US-09-610-417-2
; Sequence 2, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:15:52 ; Search time 23 Seconds
(without alignments)
1263.714 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903

Sequence: 1 MARKQNSKELGLVPLTDD.....NTAWIQAATTTATKEHPL 563

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pdp.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pdp.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pdp.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pdp.*
- 5: /cgm2_6/ptodata/2/iaa/PCTUS.COMB.pdp.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2903	100.0	563	2	US-09-031-392-2
2	2903	100.0	563	3	US-09-299-549-2
3	2903	100.0	563	4	US-09-610-417-2
4	1108	38.2	500	2	US-09-031-392-7
5	1108	38.2	500	3	US-09-299-549-7
6	1108	38.2	500	4	US-09-610-417-7
7	1081	37.2	493	2	US-09-031-392-10
8	1081	37.2	493	3	US-09-299-549-10
9	1081	37.2	493	4	US-09-610-417-10
10	919.5	31.7	492	2	US-08-355-844-3
11	919.5	31.7	492	5	PCT-US95-16126-3
12	874	30.1	509	2	US-09-031-392-6
13	874	30.1	509	3	US-09-299-549-6
14	874	30.1	509	4	US-09-610-417-6
15	866	29.8	494	2	US-09-031-392-5
16	866	29.8	494	3	US-09-299-549-5
17	866	29.8	494	4	US-09-610-417-5
18	823.5	28.4	863	4	US-09-894-927B-9
19	751	25.9	534	2	US-09-031-392-4
20	751	25.9	534	3	US-09-299-549-4
21	751	25.9	534	4	US-09-610-417-4
22	725.5	25.0	383	2	US-09-031-392-3
23	725.5	25.0	383	3	US-09-299-549-3
24	725.5	25.0	383	4	US-09-610-417-3
25	695.5	24.0	524	2	US-08-928-692-12
26	695.5	24.0	524	4	US-09-339-972-12
27	600	20.7	488	4	US-10-162-012-46

28	480	16.5	488	2	US-08-928-692-11	Sequence 11, Appl
29	480	16.5	488	4	US-09-339-972-11	Sequence 11, Appl
30	422.5	14.6	549	4	US-09-291-922-30	Sequence 30, Appl
31	419	14.4	523	4	US-09-291-922-24	Sequence 24, Appl
32	418.5	14.4	476	4	US-09-489-039A-11933	Sequence 11933, A
33	417.5	14.4	501	4	US-09-489-039A-11731	Sequence 11731, A
34	397.5	13.7	510	4	US-09-291-922-22	Sequence 22, Appl
35	396.5	13.7	502	4	US-09-679-686B-2	Sequence 2, Appl
36	390	13.4	529	4	US-09-291-922-28	Sequence 28, Appl
37	389.5	13.4	510	4	US-09-679-686B-19	Sequence 19, Appl
38	389	13.4	514	4	US-09-489-039A-11902	Sequence 11902, A
39	382.5	13.2	488	2	US-08-928-692-10	Sequence 10, Appl
40	382.5	13.2	488	4	US-09-339-972-10	Sequence 10, Appl
41	382	13.2	514	4	US-09-679-686B-22	Sequence 22, Appl
42	374.5	12.9	584	2	US-08-928-692-13	Sequence 13, Appl
43	374.5	12.9	584	4	US-09-339-972-13	Sequence 13, Appl
44	372	12.8	539	4	US-09-291-922-26	Sequence 26, Appl
45	369	12.7	518	4	US-09-679-686B-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-031-392-2
; Sequence 2, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 563 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-031-392-2

Query Match 100.0%; Score 2903; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 3.9e-280;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MARKQNSKELGLVPLTDDTSHAGPPGCRALLECDHLRSVPGGRRRDKWSCSLIVAS 60
Db 1 MARKQNSKELGLVPLTDDTSHAGPPGCRALLECDHLRSVPGGRRRDKWSCSLIVAS 60

QY 347 GIPPAKIPVVTLSGTGIEETLAAVFGSLVTEHIGRPLLIIGGFLMGLPFGTITITLQD 406
Db 318 GIPLAKIPVVTLSGTGIEETLAAVFGSLVTEHIGRPLLIIGGFLMGLPFGTITITLQD 377
QY 407 HAPWVYLSIVGILALIIASFCSGPGGIP 434
Db 378 HAPWVYLSIVGILALIIASFCSGPAVFP 405
Search completed: April 6, 2004, 10:16:58
Job time : 62 secs

XX 22-DEC-2000; 2000WO-US035095.
PF 23-DEC-1999; 99US-0172000P.
XX 14-JAN-2000; 2000US-0176083P.
PR 21-JAN-2000; 2000US-0177332P.
PR 28-JAN-2000; 2000US-0178572P.
PR 02-FEB-2000; 2000US-0179758P.
PR 10-FEB-2000; 2000US-0181625P.
XX (INCY-) INCYTE GENOMICS INC.
XX Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R;
PI Lal P, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
PI Tang YT, Khan FA;
XX WPI; 2001-418042/44.
DR N-PSDB; AAD09557.
XX Novel human transporter and ion channel proteins useful for treating and
PT preventing transport, neurological, muscle and immunological disorders.
PT Claim 1; Page 118-119; 160pp; English.
PS
XX The present sequence is transporter and ion channel-6 (TRICH-6) protein.
CC TRICH is used as vaccine. TRICH is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH, such
CC as transport disorder including amyotrophic lateral sclerosis, cystic
CC fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease,
CC Duchenne muscular dystrophy, angina and hypertension, neurological
CC disorders including Alzheimer's disease, amnesia, bipolar disorder,
CC dementia, depression, epilepsy, ischaemic cerebrovascular disease,
CC stroke, cerebral neoplasms, Pick's disease, Huntington's disease and
CC Parkinson's disease, demyelinating diseases, mental disorders including
CC mood, anxiety, Schizophrenia and seasonal affective disorder, muscle
CC disorder including cardiomyopathy, myocarditis, polymyositis,
CC dermatomyositis, arrhythmias and asthma and immunological disorders
CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC syndrome, systemic lupus erythematosus and other diseases including
CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
CC infections. TRICH DNA is useful in gene therapy and in diagnostic
CC purposes
XX Sequence 416 AA;
SQ

Query Match 66.8%; Score 1939; DB 4; Length 416;
Best Local Similarity 97.9%; Pred. No. 9.6e-189;
Matches 380; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 47 RRRKDWSCSLVASLAGAFSGSFLYGYNLVSVNAPTPIKAFYNESWERRHGRIPDPTL 106
Db 18 KKKLDWSCSLVASLAGAFSGSFLYGYNLVSVNAPTPIKAFYNESWERRHGRIPDPTL 77
QY 107 TLLMSVTYSIFAIGGLVGTLLVFMKIGVKLGRKHTLLANNPFAISAALLMACSLQAGAFEM 166
Db 78 TLLMSVTYSIFAIGGLVGTLLVFMKIGVKLGRKHTLLANNPFAISAALLMACSLQAGAFEM 137
QY 167 LIVGRFIMGIDGGVALSVLPHYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGPPELLG 226
Db 138 LIVGRFIMGIDGGVALSVLPHYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGPPELLG 197
QY 227 KESTWPYLFVGVVPAVVQVLLSLPFLPDSPPYLLLEKHENARAVKAFOTFLGKADVSQEV 286
Db 198 KESTWPYLFVGVVPAVVQVLLSLPFLPDSPPYLLLEKHENARAVKAFOTFLGKADVSQEV 257
QY 287 BEVLAESHVQSRIRLVSVLEILRAPYVRWQVVTIVTMACYQLCGLNAINFYTNSIFGKA 346
Db 258 BEVLAESHVQSRIRLVSVLEILRAPYVRWQVVTIVTMACYQLCGLNAINFYTNSIFGKA 317

Best Local Similarity 99.5%; Pred. No. 3.3e-196; Mismatches 2; Indels 0; Gaps 0; Matches 395; Conservative 0

QY 137 RKHTLLANNPFAISALLMACSLQAGAFEMLVGRFIMGIDGVALSVLPMLSEISPK 196
DB 5 RKHTLLANNPFAISALLMACSLQAGAFEMLVGRFIMGIDGVALSVLPMLSEISPK 64

QY 197 IRGSLGQVTAIFTCIGVFTGQLLGLPELKGESTWPLFGVIVPAVQVLLSLPLPDPSP 256
DB 65 IRGSLGQVTAIFTCIGVFTGQLLGLPELKGESTWPLFGVIVPAVQVLLSLPLPDPSP 124

QY 257 RYLLLEKNEARAVKAPQTFPLGKADVSQVEVEVLAESHVQSRIRLVLELLRAPVVRWQ 316
DB 125 RYLLLEKNEARAVKAPQTFPLGKADVSQVEVEVLAESHVQSRIRLVLELLRAPVVRWQ 184

QY 317 VVVIVTMACYOICGLNAINFYTNSIFGKAGIPPAKIPVVTLSGTGIELTAAVFSGLVIE 376
DB 185 VVVIVTMACYOICGLNAINFYTNSIFGKAGIPPAKIPVVTLSGTGIELTAAVFSGLVIE 244

QY 377 HLGRRPLLIGFGMLGFFGTITLTLODHPWPVLSIVGILAIASFCSGGGIPFI 436
DB 245 HLGRRPLLIGFGMLGFFGTITLTLODHPWPVLSIVGILAIASFCSGGGIPFI 304

QY 437 LTGFEFFQSQRPAAFIAGTVNMLSNFAVGLLPFFIOKSLDTYCFVLFATICTGAIYLY 496
DB 305 LTGFEFFQSQRPAAFIAGTVNMLSNFAVGLLPFFIOKSLDTYCFVLFATICTGAIYLY 364

QY 497 FVLPEKNTVABEISQAFSKRNKAYPPEEKIDSAVTD 533
DB 365 FVLPEKNTVABEISQAFSKRNKAYPPEEKIDSAVTD 401

RESULT 15
AAE04893
ID AAE04893 standard; protein; 416 AA.
XX AAE04893;
AC AAE04893;
XX 10-SEP-2001 (first entry)
DT
XX Human transporter and ion channel-6 (TRICH-6) protein.
DE
XX Human; transporter and ion channel-6; TRICH-6; vaccine; cystic fibrosis; gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia; Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasia; allergy; demyelinating disease; mental disorder; Schizophrenia; polymyositis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer.
XX Homo sapiens.
OS
XX Key
FH Peptide
FT Peptide
FT Domain
FT Domain
FT Protein
FT Domain
FT Domain
FT Domain
FT Domain
XX WO200146258-A2.
XX 28-JUN-2001.
PD

QY 497 FVLPEKNTVABEISQAFSKRNKAYPPEEKIDSAVTD 533
DB 371 FVLPEKNTVABEISQAFSKRNKAYPPEEKIDSAVTD 407

RESULT 14
AAM39530
ID AAM39530 standard; protein; 408 AA.
XX AAM39530;
AC AAM39530;
XX 22-OCT-2001 (first entry)
DT
XX Human polypeptide SEQ ID NO 2675.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX
OS Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000NO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AAI59686.
XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
XX Example 4; SEQ ID NO 2675; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM4213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification
XX Sequence 408 AA;
SQ

Query Match 69.3%; Score 2012; DB 4; Length 408;

XX SQ Sequence 408 AA; Query Match 69.5%; Score 2018; DB 4; Length 408; Best Local Similarity 99.5%; Pred. No. 8e-197; Mismatches 1; Indels 0; Gaps 0; Matches 395; Conservative

QY 137 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 196
DB 5 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 64
QY 197 IRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYPLFGVIVVPAVQVLLSLPFLPDSP 256
DB 65 IRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYPLFGVIVVPAVQVLLSLPFLPDSP 124
QY 257 RYLLEKHNREARAVKAFQTFGLKADVSQEEVEVLAEASHVQSRIRLVSVLELLRAPYVRWQ 316
DB 125 RYLLEKHNREARAVKAFQTFGLKADVSQEEVEVLAEASHVQSRIRLVSVLELLRAPYVRWQ 184
QY 317 VVTIVVTMACYQLCGLNAINWFYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 376
DB 185 VVTIVVTMACYQLCGLNAINWFYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 244
QY 377 HLGRRPELLIGGFGMLGFLFPGTTLITLTLDHAPWPVYLSIVGILAIIASFCSPGGGIPPI 436
DB 245 HLGRRPELLIGGFGMLGFLFPGTTLITLTLDHAPWPVYLSIVGILAIIASFCSPGGGIPPI 304
QY 437 LTGEFFQOSORPAAFIAGTWNLSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 496
DB 305 LTGEFFQOSORPAAFIAGTWNLSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 364
QY 497 FVLPEKTRNRYVAEISQAFSKENKAYPPEEKIDSAVTD 533
DB 365 FVLPEKTRNRYVAEISQAFSKENKAYPPEEKIDSAVTD 401

RESULT 13
ID AD26265 standard; protein; 422 AA.
XX AC AD26265;
XX AC AD26265;
XX DT 18-DEC-2003 (first entry)
XX DE Human NOV17h protein.
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
XX KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
XX KW transgenic; human.
XX OS Homo sapiens.
XX OS WO2003004697-A2.
XX PD 16-JAN-2003.
XX PF 03-JUL-2002; 2002WO-US021361.
XX PR 05-JUL-2001; 2001US-0303046P.
XX PR 09-JUL-2001; 2001US-0303828P.
XX PR 09-JUL-2001; 2001US-0304016P.
XX PR 11-JUL-2001; 2001US-0304502P.
XX PR 13-JUL-2001; 2001US-0305262P.
XX PR 16-JUL-2001; 2001US-0305673P.
XX PR 17-JUL-2001; 2001US-0306085P.
XX PR 24-JUL-2001; 2001US-0307536P.
XX PR 27-JUL-2001; 2001US-0308228P.
XX PR 30-JUL-2001; 2001US-0308877P.
XX PR 01-AUG-2001; 2001US-0309255P.
XX PR 17-AUG-2001; 2001US-0313328P.
XX PR 12-SEP-2001; 2001US-0318711P.
XX PR 19-SEP-2001; 2001US-0323380P.
XX PR 21-SEP-2001; 2001US-0323969P.

PR 04-JAN-2002; 2002US-0345022P.
PR 04-JAN-2002; 2002US-0345038P.
PR 28-FEB-2002; 2002US-0361172P.
PR 01-MAR-2002; 2002US-0360814P.
PR 01-MAR-2002; 2002US-0360830P.
PR 01-MAR-2002; 2002US-0361133P.
PR 01-MAR-2002; 2002US-0361147P.
PR 05-MAR-2002; 2002US-0361677P.
PR 02-APR-2002; 2002US-0363637P.
PR 12-APR-2002; 2002US-0372326P.
PR 16-APR-2002; 2002US-0372990P.
PR 19-APR-2002; 2002US-0373881P.
PR 19-APR-2002; 2002US-0373921P.
PR 02-JUL-2002; 2002US-00188186.
XX (CURA-) CUBAGEN CORP.
XX PA Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ, Catterton B, Edinger S, Eissen AU, Ellerman K, Gerlach V, Gorman L, Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE, Padigaru M, Patturajan M, Pena CEA, Rastelli L, Shenoy S, Shinkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
XX WPI; 2003-221607/21.
XX N-PSDB; AD26264.
XX New isolated NOVX polypeptide, useful for determining the presence of, or predisposition to a disease associated with altered levels of expression of the polypeptide, and for treating or preventing cancer.
XX Claim 2; SEQ ID NO 90; 478pp; English.
XX The invention relates to a novel isolated NOV polypeptide. The polypeptide of the invention demonstrates cytotostatic activity and may be used for determining the presence of, or predisposition to a disease associated with altered levels of expression of the polypeptide, including metabolic disorders, immune disorders, neurodegenerative disorders, circulatory diseases, haemopoietic disorders, wasting diseases and cancer. The polypeptide may also be utilised during gene therapy procedures, vaccine development and transgenic animal production. The current sequence is that of the human NOV protein of the invention.

Query Match 69.5%; Score 2017; DB 7; Length 422;
Best Local Similarity 99.5%; Pred. No. 1.1e-196;
Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 137 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 196
DB 11 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 70
QY 197 IRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYPLFGVIVVPAVQVLLSLPFLPDSP 256
DB 71 IRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYPLFGVIVVPAVQVLLSLPFLPDSP 130
QY 257 RYLLEKHNREARAVKAFQTFGLKADVSQEEVEVLAEASHVQSRIRLVSVLELLRAPYVRWQ 316
DB 131 RYLLEKHNREARAVKAFQTFGLKADVSQEEVEVLAEASHVQSRIRLVSVLELLRAPYVRWQ 190
QY 317 VVTIVVTMACYQLCGLNAINWFYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 376
DB 191 VVTIVVTMACYQLCGLNAINWFYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 250
QY 377 HLGRRPELLIGGFGMLGFLFPGTTLITLTLDHAPWPVYLSIVGILAIIASFCSPGGGIPPI 436
DB 251 HLGRRPELLIGGFGMLGFLFPGTTLITLTLDHAPWPVYLSIVGILAIIASFCSPGGGIPPI 310
QY 437 LTGEFFQOSORPAAFIAGTWNLSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 496
DB 311 LTGEFFQOSORPAAFIAGTWNLSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 370

PR 05-MAR-2002; 2002US-0361677P.
 PR 02-APR-2002; 2002US-0363637P.
 PR 12-APR-2002; 2002US-0372326P.
 PR 16-APR-2002; 2002US-0372990P.
 PR 19-APR-2002; 2002US-0372981P.
 PR 19-APR-2002; 2002US-0373881P.
 PR 02-JUL-2002; 2002US-00188186.
 XX (CURA-) CURAGEN CORP.
 PA Anderson DW, Berghs C, Boidog FL, Burgess CE, Casman SJ, Catterson E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L, Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE, Radigaru M, Patturajan M, Pena CE, Raetelli L, Shenoy S, Shimketa RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
 XX WPI: 2003-221607/21.
 DR N-PSDB; ADC26252.
 XX
 PT New isolated NOVX polypeptide, useful for determining the presence of, or
 PT predisposition to a disease associated with altered levels of expression
 PT of the polypeptide, and for treating or preventing cancer.
 XX
 PS Claim 2; SEQ ID NO 78; 478pp; English.
 XX
 CC The invention relates to a novel isolated NOV polypeptide. The
 CC polypeptide of the invention demonstrates cytostatic activity and may be
 CC used for determining the presence of, or predisposition to a disease
 CC associated with altered levels of expression of the polypeptide,
 CC including metabolic disorders, immune disorders, neurodegenerative
 CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases
 CC and cancer. The polypeptide may also be utilised during gene therapy
 CC procedures, vaccine development and transgenic animal production. The
 CC current sequence is that of the human NOV protein of the invention.
 XX
 XX Sequence 565 AA;
 SQ
 Query Match 80.5%; Score 2337.5; DB 7; Length 565;
 Best Local Similarity 77.3%; Pred. No. 3e-229;
 Matches 474; Conservative 0; Mismatches 4; Indels 135; Gaps 2;
 QY 1 MARKQNSKELGLVPLTDDTSHAGPPGPGGRALLECDHLRSVPGRRK----- 50
 DB 1 MARKQNSKELGLVPLTDDTSHAGPPGPGGRALLECDHLRSVPGRRKQPLRSTSSAA 60
 QY 51 ----- 50
 DB 61 GSSTTYVASAAKIPFHRWASLAPEKSHLLRLQLGLALEVKLISLGSKPOSSISWL 120
 QY 51 -----DWSCLLVASLAGAFSGFLYGNLSVYNAPTPIYKAFYNESWERRHGRP 100
 DB 121 ALEETNEIGRDSCLLVASLAGAFSGFLYGNLSVYNAPT----- 163
 QY 101 IDPDTLLMSVTVSIFAIGLVGTLLIVKMGKVLGRKHLLANNGPAISAALLMACSLQ 160
 DB 164 -----HTLLANNGPAISAALLMACSLQ 185
 QY 161 AGAFEMLIIVGRFIMGIDGGVALSVLPMVLSEISPEKRGISQVTAIFIGVFTGQLG 220
 DB 186 AGAFEMLIIVGRFIMGIDGGVALSVLPMVLSEISPEKRGISQVTAIFIGVFTGQLG 245
 QY 221 LPELLGKESTWYLFVGVIVPAVQVLLSLPLPDSPRYLLLEKNEARAVKAFQTFLGKA 280
 DB 246 LPELLGKESTWYLFVGVIVPAVQVLLSLPLPDSPRYLLLEKNEARAVKAFQTFLGKA 305
 QY 281 DVSQEVEVLAESHVQSRIRLVSVLELRAPVVRQVTVTVWVACVQLCGLNAIWYTN 340
 DB 306 DVSQEVEVLAESHVQSRIRLVSVLELRAPVVRQVTVTVWVACVQLCGLNAIWYTN 365
 QY 341 SIFGKAGIPPAKIPVYTLSTGGIETLAIVFSGLVIEHGRPLLIIGFGLMGLFPGLTI 400
 DB 366 SIFGKAGIPPAKIPVYTLSTGGIETLAIVFSGLVIEHGRPLLIIGFGLMGLFPGLTI 425

QY 401 TLTLDHAPWVYLSIVGILAIASFCSPGGIPFILTGEFFQOSORPAAFIAGTVNWL 460
 DB 426 TLTLDHAPWVYLSIVGILAIASFCSPGGIPFILTGEFFQOSORPAAFIAGTVNWL 485
 QY 461 SNFAVGLLFPFIQKSLDTCFLVFAFATCITGALYLYFLVLPETKNTYAEISQAFSKENKA 520
 DB 486 SNFAVGLLFPFIQKSLDTCFLVFAFATCITGALYLYFLVLPETKNTYAEISQAFSKENKA 545
 QY 521 YPPEEKIDSAVTD 533
 DB 546 YPPEEKIDSAVTD 558
 XX AAM78438 standard; protein; 408 AA.
 AC AAM78438;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1100.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00820325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR N-PSDB; AAK51571.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 3326-3327; 622pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIYGRFIMGIDGV 180
121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIYGRFIMGIDGV 180
181 ALSVLPWYLSSEISPKIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYLFVIVV 240
181 ALSVLPWYLSSEISPKIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKSTWYLFVIVV 240
241 PAVVQLLSLPFLPDSRYLLLEKHNEARAVKAFQTFELGKADYSQEVVEVLAESHVORSIR 300
241 PAVVQLLSLPFLDSDRYLLLEKHNEARAVKAFQTFELGKADYSQEVVEVLAESHVORSIR 300
301 LVSLELLRAPYVRWQVTVITMACYQLCGLNAIFWYNSIFGKAGIPAKIPYTLST 360
301 LVSLELLRAPYVRWQVTVITMACYQLCGLNAIFWYNSIFGKAGIPAKIPYTLST 360
361 GGIETLAAPVSGLVIEHLGRPLLIGGFLMGLFFGTLTITLTLQDHAPWVPLSIVGIL 420
361 GGIETLAAPVSGLVIEHLGRPLLIGGFLMGLFFGTLTITLTLQDHAPWVPLSIVGIL 420
421 AIIASFCSGGGIPFILTGTGEFFOOSORPAAFIAGTWNLSNPAVGLLPPFIKSLDTYC 480
421 AIIASFCSGGGIPFILTGTGEFFOOSORPAAFIAGTWNLSNPAVGLLPPFIKSLDTYC 480
372 -----GIPFILTGTGEFFOOSORPAAFIAGTWNLSNPAVGLLPPFIKSLDTYC 420
481 FLVFAITCITGAIYLVFLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
481 FLVFAITCITGAIYLVFLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
transgenic; human.
Homo sapiens.
WO2003004687-A2.
16-JAN-2003.
03-JUL-2002; 2002WO-US021361.
05-JUL-2001; 2001US-0303046P.
09-JUL-2001; 2001US-0303828P.
09-JUL-2001; 2001US-0304016P.
11-JUL-2001; 2001US-0304502P.
13-JUL-2001; 2001US-0305262P.
16-JUL-2001; 2001US-0305673P.
17-JUL-2001; 2001US-0306085P.
24-JUL-2001; 2001US-0307536P.
27-JUL-2001; 2001US-0308228P.
30-JUL-2001; 2001US-0308877P.
01-AUG-2001; 2001US-0309255P.
12-SEP-2001; 2001US-0313328P.
12-SEP-2001; 2001US-0318711P.
19-SEP-2001; 2001US-0323380P.
21-SEP-2001; 2001US-0323969P.
04-JAN-2002; 2002US-0345022P.
04-JAN-2002; 2002US-0345038P.
28-FEB-2002; 2002US-0361172P.
01-MAR-2002; 2002US-0360814P.
01-MAR-2002; 2002US-0360830P.
01-MAR-2002; 2002US-0361133P.
01-MAR-2002; 2002US-0361147P.

05-JUL-2001; 2001US-0303046P.
09-JUL-2001; 2001US-0303828P.
09-JUL-2001; 2001US-0304016P.
11-JUL-2001; 2001US-0304502P.
13-JUL-2001; 2001US-0305262P.
16-JUL-2001; 2001US-0305673P.
17-JUL-2001; 2001US-0306085P.
24-JUL-2001; 2001US-0307536P.
27-JUL-2001; 2001US-0308228P.
30-JUL-2001; 2001US-0308877P.
01-AUG-2001; 2001US-0309255P.
17-AUG-2001; 2001US-0313328P.
12-SEP-2001; 2001US-0318711P.
19-SEP-2001; 2001US-0323380P.
21-SEP-2001; 2001US-0323969P.
04-JAN-2002; 2002US-0345022P.
04-JAN-2002; 2002US-0345038P.
28-FEB-2002; 2002US-0361172P.
01-MAR-2002; 2002US-0360814P.
01-MAR-2002; 2002US-0360830P.
01-MAR-2002; 2002US-0361133P.
01-MAR-2002; 2002US-0361147P.
05-MAR-2002; 2002US-0361677P.
02-APR-2002; 2002US-0363637P.
12-APR-2002; 2002US-0372326P.
16-APR-2002; 2002US-0372990P.
19-APR-2002; 2002US-0373881P.
19-APR-2002; 2002US-0373921P.
02-JUL-2002; 2002US-00188186.
(CURA-) CURAGEN CORP.
Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ,
Catterton EJ, Edinger S, Eissen AJ, Ellerman K, Gerlach V, Gorman L,
Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE,
Padiguru M, Patturajan M, Pena CEA, Rastelli L, Shenoy S,
Shimkete RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ,
Vernet CAM, Voss EZ, Zhong M;
WPI; 2003-221607/21.
N-PSDB; ADC26254.
New isolated NOVX polypeptide, useful for determining the presence of, or
predisposition to a disease associated with altered levels of expression
of the polypeptide, and for treating or preventing cancer.
Claim 2; SEQ ID NO 80; 478pp; English.
The invention relates to a novel isolated NOV polypeptide. The
polypeptide of the invention demonstrates cytostatic activity and may be
used for determining the presence of, or predisposition to a disease
associated with altered levels of expression of the polypeptide,
including metabolic disorders, immune disorders, neurodegenerative
disorders, circulatory diseases, haemopoietic disorders, wasting diseases
and cancer. The polypeptide may also be utilised during gene therapy
procedures, vaccine development and transgenic animal production. The
current sequence is that of the human NOV protein of the invention.
Sequence 480 AA;
Query Match 80.6%; Score 2341; DB 7; Length 480;
Best Local Similarity 87.6%; Pred. No. 1e-229; 6; Indels 60; Gaps 1;
Matches 467; Conservative 0; Mismatches 6;
1 MARKQNRNSELGLVPLTDDTSHAGPPGFRALLECDHLRSGVPGRRKDWSCSLVAS 60
1 MARKQNRNSELGLVPLTDDTSHAGPPGFRALLECDHLRSGVPGRRKDWSCSLVAS 60
61 LAGAFSGFLYGNLSVNNAPPTYKAFYNESWERRHGRPIDPDLTLLWSVTVSIFAIG 120
61 LAGAFSGFLYGNLSVNNAPPTYKAFYNESWERRHGRPIDPDLTLLWSVTVSIFAIG 120

DR N-PSDB; AA160472.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Example 2; SEQ ID NO 6247; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the

CC encoded polypeptides (AA158642-AA162213) with neurotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX Seq Sequence 558 AA;

Query Match 94.0%; Score 2730; DB 4; Length 558;

Best Local Similarity 99.6%; Pred. No. 2.4e-269;

Matches 53; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MARKQNRNSKELGLVPLTDDTSAGPPGPRALLECDHLRSYGPGRRKQWCSLLVAS 60

Db 19 MARKQNRNSKELGLVPLTDDTSAGPPGPRALLECDHLRSYGPGRRKQWCSLLVAS 78

Qy 61 LAGNFGSSFLYGNLSVNNAPTPVIKAFYNESWERRHGRPIDDTLTLLMSVTSIFAIG 120

Db 79 LAGAFSSFLYGNLSVNNAPTPVIKAFYNESWERRHGRPIDDTLTLLMSVTSIFAIG 138

Qy 121 GLVGTLLVMKIGKVLGRKHTLLANNFPAISAAALLMACSLQAGAFEMLI VGRFIMGIDGV 180

Db 139 GLVGTLLVMKIGKVLGRKHTLLANNFPAISAAALLMACSLQAGAFEMLI VGRFIMGIDGV 198

Qy 181 ALSVLPYMLSEISPEKIRGSLQVTAIFICIGVFTGQLGLPBLLEGSTWPLFGVIV 240

Db 199 ALSVLPYMLSEISPEKIRGSLQVTAIFICIGVFTGQLGLPBLLEGSTWPLFGVIV 258

Qy 241 PAVVQLLSPLPDSRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEASHVQSR 300

Db 259 PAVVQLLSPLPDSRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEASHVQSR 318

Qy 301 LVSVELLRAPYVRQVTVVTVWACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

Db 319 LVSVELLRAPYVRQVTVVTVWACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 378

Qy 361 GGIELLAIVFSGLVIEHLGRPLLIIGFGLMGLFPGTLITLTLDHAPWPVYLSIVGIL 420

Db 379 GGIELLAIVFSGLVIEHLGRPLLIIGFGLMGLFPGTLITLTLDHAPWPVYLSIVGIL 438

Qy 421 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480

Db 439 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 498

Qy 481 FLVFATICTGAIVLYFVLPETKRTYAEISQAFSKRNKAYPPEEKIDSAVTD 533

Db 499 FLVFATICTGAIVLYFVLPETKRTYAEISQAFSKRNKAYPPEEKIDSAVTD 551

RESULT 7

AAAE16788

ID AAE16788 standard; protein; 537 AA.

XX AAE16788;

AC AAE16788;

XX 09-APR-2002 (first entry)

DT

XX Human transporter and ion channel-25 (TRICH-25) protein.

XX Human; transporter and ion channel-25; TRICH-25; neuroprotective; asthma;

XX neurotropic; cytostatic; cardiovascular; immunosuppressive; cardiomyopathy;

KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;

KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;

KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;

KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;

KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;

KW bradyarrhythmia; gene expression; drug screening.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..66

XX Domain /label= Signal_peptide

XX Protein /note= "Sugar transporter domain"

XX Domain 59..514

XX Domain 67..537

XX Domain 112..128

XX Domain /label= Mature_TRICH_25_protein

XX Domain 173..198

XX Domain /label= Transmembrane_domain

XX Domain /note= "Sugar transporter 2 motif"

XX Domain 371..386

XX Domain /note= "Sugar transporter 1 motif"

XX Domain 385..404

XX Domain /label= Transmembrane_domain

XX Domain 414..436

XX Domain /label= Transmembrane_domain

XX Domain 479..497

XX Domain /label= Transmembrane_domain

XX WO200192304-A2.

XX 06-DEC-2001.

XX 25-MAY-2001; 2001WO-US017065.

XX 26-MAY-2000; 2000US-0208424P.

XX 01-JUN-2000; 2000US-0209001P.

XX 08-JUN-2000; 2000US-0210588P.

XX 16-JUN-2000; 2000US-0212335P.

XX 22-JUN-2000; 2000US-0213747P.

XX 29-JUN-2000; 2000US-0215391P.

XX (INCY-) INCYTE GENOMICS INC.

XX Thornton M, Wallia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;

PI Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang Y;

PI Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BE;

PI Lee EA, Hafalia A, Baughn MR, Green BD, Khan FA, Kearney L;

PI Elliot VS, Seilhamer JJ, Policky JL, Borowsky ML, Burford N, Ding L;

PI Lu DAM, Hillman JL;

XX WPI; 2002-122055/16.

XX N-PSDB; AAD27278.

XX New human transporters and ion channels (TRICH) polypeptides useful for

PT diagnosing, treating or preventing disorders associated with aberrant

XX expression of TRICH.

XX Claim 1; Page 182-184; 210pp; English.

XX The invention relates to human transporters and ion channels (TRICH)

CC polypeptides and their cDNA molecules. The nucleic acid and polypeptide

CC sequences are useful in the diagnosis, treatment, and prevention of

CC disorders associated with transport (akinesia, cystic fibrosis, Bell's

CC palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,

CC amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's

CC muscular dystrophy); immunological (AIDS, Addison's disease, allergies;

CC asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);

Db 79 LAGAFSGSFLYGNLSVWVNAPTPIYKAFYNESWERRHGRPIDPDTLTLLWSVTVSIFATG 138
QY 121 GLVGTLLIVKMGVGLGRKHTLLANNNGFAJSAALLMACSLQAGAFEMLIYGRFTMGIDGV 180
Db 139 GLVGTLLIVKMGVGLGRKHTLLANNNGFAJSAALLMACSLQAGAFEMLIYGRFTMGIDGV 198
QY 181 ALSVLPMYLSEISPKIRSGSLQVTAIFICIGVFTQQLLGLPELLGKSTWPLFQVIVV 240
Db 199 ALSVLPMYLSEISPKIRSGSLQVTAIFICIGVFTQQLLGLPELLGKSTWPLFQVIVV 258
QY 241 PAVVQLLSLPPFLDPSRYLLEKHNARAVKAPQTFGLKADVSQEVVEVLAESHVORSR 300
Db 259 PAVVQLLSLPPFLDPSRYLLEKHNARAVKAPQTFGLKADVSQEVVEVLAESHVORSR 318
QY 301 LVSVELELLRAPYRWQVTVITMACYQVLCGLNAINFYTNISFGKAGIPPAKIPYVTLST 360
Db 319 LVSVELELLRAPYRWQVTVITMACYQVLCGLNAINFYTNISFGKAGIPPAKIPYVTLST 378
QY 361 GGIETLAAVFSGLVIEHLGRRLIIGFGLMGLFFGTTLITTLQDHAPWVPLSIVGIL 420
Db 379 GGIETLAAVFSGLVIEHLGRRLIIGFGLMGLFFGTTLITTLQDHAPWVPLSIVGIL 438
QY 421 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNMLSNFAVGLLPPFIQKSLDTYC 480
Db 439 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNMLSNFAVGLLPPFIQKSLDTYC 498
QY 481 FLVPATICITGAIYLYFVLPETKQRYAEISQAFSKRNKAYPPEEKIDSAVTD 533
Db 499 FLVPATICITGAIYLYFVLPETKQRYAEISQAFSKRNKAYPPEEKIDSAVTD 551
RESULT 6
AAM41316
ID AAM41316 standard; protein; 558 AA.
XX AC AAM41316;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 6247.
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR

RESULT 5
AAM79422
ID AAM79422 standard; protein; 558 AA.
XX AC AAM79422;
XX DT 06-NOV-2001 (first entry)
XX DE Human protein SEQ ID NO 3068.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US004098.
XX PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX N-PSDB; AAK52555.
XX KW Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX PS Claim 20; Page 242; 6221pp; English.
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX SQ Sequence 558 AA;
Query Match 94.3%; Score 2737; DB 4; Length 558;
Best Local Similarity 99.8%; Pred. No. 4.7e-270;
Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSGVPGRRKDKWSCSILLVAS 60
Db 19 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSGVPGRRKDKWSCSILLVAS 78
QY 61 LAGAFSGSFLYGNLSVWVNAPTPIYKAFYNESWERRHGRPIDPDTLTLLWSVTVSIFATG 120

Query Match	94.3%;	Score 2737;	DB 4;	Length 558;
Best Local Similarity	99.8%;	Pred. No. 4.7e-270;		
Matches 532;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
QY	1	MARKQNRNSKELGLVPLTDDTSHAGPPGPGHALLCECHLRSGVGGRRKDWSCSLVAS	60	
DB	19	MARKQNRNSKELGLVPLTDDTSHAGPPGPGHALLCECHLRSGVGGRRKDWSCSLVAS	78	
QY	61	LAGAFGSSFLGYCNLSVNNAPTYIKAFYNESWERRHGRPDPTLTLLWSVTYSIFAIG	120	
DB	79	LAGAFGSSFLGYCNLSVNNAPTYIKAFYNESWERRHGRPDPTLTLLWSVTYSIFAIG	138	
QY	121	GLVGTLLVKMTGKVLGRKHTLLANGPAISAAALLMACSQAGAFEMLVGSRFIMGIDGGV	180	
DB	139	GLVGTLLVKMTGKVLGRKHTLLANGPAISAAALLMACSQAGAFEMLVGSRFIMGIDGGV	198	
QY	181	ALSVLPMYLSISPEIKRISGLGVQVTAIFICIGVFTGQLLGLPELLGKESTWPYLFGVTV	240	
DB	199	ALSVLPMYLSISPEIKRISGLGVQVTAIFICIGVFTGQLLGLPELLGKESTWPYLFGVTV	258	
QY	241	PAVQLLSLPFDPSPRYLLLEKNEARAVKAFOTFLGKADVSEQVEEVLAEASHVQSR	300	
DB	259	PAVQLLSLPFDPSPRYLLLEKNEARAVKAFOTFLGKADVSEQVEEVLAEASHVQSR	318	
QY	301	LVSVLELLRAPVVRQWVTVIVTMACYOGLNAINFWYTNISFGKAGIPPAKIPVYTLST	360	
DB	319	LVSVLELLRAPVVRQWVTVIVTMACYOGLNAINFWYTNISFGKAGIPPAKIPVYTLST	378	
QY	361	GGIETLAAVFGSLVIEHLGRPLLIGFGLMGLPFGTLTITLTLODHAPWPVYLSIVGIL	420	
DB	379	GGIETLAAVFGSLVIEHLGRPLLIGFGLMGLPFGTLTITLTLODHAPWPVYLSIVGIL	438	
QY	421	AIIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWLSNFAVGLLPPFFTKSLDTC	480	
DB	439	AIIASFCSPGGIPFILTGEFFQSQORPAAFIAGTVNWLSNFAVGLLPPFFTKSLDTC	498	
QY	481	FLVFPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD	533	
DB	499	FLVFPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD	551	

CC specific markers
XX
SQ Sequence 563 AA;

Query Match 100.0%; Score 2903; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 5.3e-287;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
QY 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
DB 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
DB 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
DB 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
QY 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
DB 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
QY 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360
DB 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

/label= Transmembrane_domain
233..252
/label= Transmembrane_domain
317..333
/label= Transmembrane_domain
355..375
/label= Transmembrane_domain
383..404
/label= Transmembrane_domain
413..437
/label= Transmembrane_domain
449..472
/label= Transmembrane_domain
481..499
/label= Transmembrane_domain

US6346374-B1.
12-FEB-2002..
05-JUL-2000; 2000US-00610417.
26-FEB-1998; 98US-00031392.
26-APR-1999; 99US-00299549.
(MILL-) MILLENNIUM PHARM INC.
Tartaglia LA, Weng X;
WPI; 2002-224989/28.
N-PSDB; AAD27918.

Identifying compound modulating activity of glucose transporter, useful for treating disorders mediated by transporter, comprises contacting polypeptide with test compound and comparing activity in presence and absence of compound.
Claim 1; Fig 2; 48pp; English.
The invention relates to identifying compounds that modulate the activity of glucose transporter (GLUTX). The compound identified is useful in the diagnosis and treatment of disorders associated with GLUTX activity and aberrant hexose transport. The present sequence is human GLUTX protein
Sequence 563 AA;

Query Match 100.0%; Score 2903; DB 5; Length 563;
Best Local Similarity 100.0%; Pred. No. 5.3e-287;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
QY 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
DB 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
DB 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
DB 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
QY 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
DB 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
QY 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360
DB 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

CC specific markers
XX
SQ Sequence 563 AA;

Query Match 100.0%; Score 2903; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 5.3e-287;
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRRDKWSCSLVAS 60
QY 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
DB 61 LAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLWSVTVSIFAIG 120
QY 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
DB 121 GLVGTLLIVKMGKVLGRKHTLLANGFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180
QY 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
DB 181 ALSVLPWYLSEISPKIRSGSLQGVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIV 240
QY 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
DB 241 PAVVQLLSLPFLPDSPRYLLLEKHNARAVKAFQTEFLGKADVSQVEEVLAEHVQSRIR 300
QY 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360
DB 301 LVSVELELLRAPYVRWQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

GGIETLAAVFGLVIEHLGRPLIIGGFLMGLFPGTITLTLLQDHAPWPVYLSIVGIL 420
GGIETLAAVFGLVIEHLGRPLIIGGFLMGLFPGTITLTLLQDHAPWPVYLSIVGIL 420
AIIASFCSGPGGIPFILTGEFFQSQSPAAFIAGTVNLSNFAVGLLFFPIQKSLDTC 480
AIIASFCSGPGGIPFILTGEFFQSQSPAAFIAGTVNLSNFAVGLLFFPIQKSLDTC 480
FLVFATCITGAIVLYFVLPTKRTYAEISQAFSKRNKAYPPEKIDSAVTDAPASSPF 540
FLVFATCITGAIVLYFVLPTKRTYAEISQAFSKRNKAYPPEKIDSAVTDAPASSPF 540
TTPNTAWTQAAATTTATKKEHPL 563
TTPNTAWTQAAATTTATKKEHPL 563

RESULT 3
AAE14611
ID AAE14611 standard; protein; 563 AA.
XX
XX AAE14611;
AC AAE14611;
XX
XX 18-JUN-2002 (first entry)
XX
XX Human glucose transporter protein GLUTX.
DE
XX Glucose transporter; GLUTX; hexose; human.
KW
XX Homo sapiens.
OS
XX
XX Key
FH Domain
FT
FT /label= Transmembrane_domain
FT 108..128
FT /label= Transmembrane_domain
FT 141..159
FT /label= Transmembrane_domain
FT 166..189
FT /label= Transmembrane_domain
FT 204..221
FT

XX (MILL-) MILLENNIUM PHARM INC.
 XX Tartaglia LA, Weng X;
 PI WPI; 1999-526192/44.
 XX N-PSDB; AA211502.
 DR A novel mammalian transporter protein and the gene encoding it, are
 XX useful in the diagnosis and treatment of disorders associated with
 PT aberrant sugar transport.
 PT
 XX Claim 1; Fig 2A-D; 48pp; English.
 XX
 XX This represents a human glucose transporter (GLUTX) protein. The protein
 CC can be expressed by standard recombinant methodology. GLUTX nucleic acids
 CC are useful as hybridization probes for detecting the presence of GLUTX
 CC DNA in a sample, useful for diagnosing conditions associated with
 CC aberrant expression levels of GLUTX. The GLUTX gene is also useful as a
 CC therapeutic agent for regulating translation of GLUTX mRNA, and for
 CC treatment of disorders associated with aberrant expression of GLUTX and
 CC aberrant hexose transport. It is useful for generating GLUTX specific
 CC antibodies, identifying agonists and antagonists of GLUTX, and
 CC identifying nucleic acids in other species encoding nucleic acids
 CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying
 CC the chromosomal location of GLUTX, and as tissue specific markers
 XX
 XX Sequence 563 AA;
 SQ

Query Match 100.0%; Score 2903; DB 2; Length 563;
 Best Local Similarity 100.0%; Pred. No. 5.3e-287;
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARKQNSKELGLVPLTDDTSHAGPPGPRALLECDHLRSGVPGRRKDWSCSLIVAS 60
 Db 1 MARKQNSKELGLVPLTDDTSHAGPPGPRALLECDHLRSGVPGRRKDWSCSLIVAS 60

Qy 61 LAGAFGSLGYNLVSVNAPTPYIKAFYNESWERRGRPIDPTLLMSVTVSIFAIG 120
 Db 61 LAGAFGSLGYNLVSVNAPTPYIKAFYNESWERRGRPIDPTLLMSVTVSIFAIG 120

Qy 121 GLVGTLLVMIGKVLGRKHTLLANGFAISAAALMACSLQAGAPMLIVGRFIMGIDGV 180
 Db 121 GLVGTLLVMIGKVLGRKHTLLANGFAISAAALMACSLQAGAPMLIVGRFIMGIDGV 180

Qy 181 ALSVLPWYLSISPKETSGSLGQVTAIFCIGVFTGQLLGLPELLGKSTWPLYGVIV 240
 Db 181 ALSVLPWYLSISPKETSGSLGQVTAIFCIGVFTGQLLGLPELLGKSTWPLYGVIV 240

Qy 241 PAVVQLLSLPDPSRYLLLEKNEARAVKAFQTFGLKADVSQVEEVLAEHVSQIR 300
 Db 241 PAVVQLLSLPDPSRYLLLEKNEARAVKAFQTFGLKADVSQVEEVLAEHVSQIR 300

Qy 301 LVSVLELLRAPYVRWQVTVITWACYQLCLNAINFYTNISIFGKAGIPPAKIPYVTLST 360
 Db 301 LVSVLELLRAPYVRWQVTVITWACYQLCLNAINFYTNISIFGKAGIPPAKIPYVTLST 360

Qy 361 GGIETLAAVFSGLVIEHGRPLLGFGMLGFFGLTITLTLQDHAPWVPLYSIVGIL 420
 Db 361 GGIETLAAVFSGLVIEHGRPLLGFGMLGFFGLTITLTLQDHAPWVPLYSIVGIL 420

Qy 421 AIIASFCSGGPILGTGEFFQSQRPAAPIIAGTVNLSNPAVGLLFPPIQKSLTYC 480
 Db 421 AIIASFCSGGPILGTGEFFQSQRPAAPIIAGTVNLSNPAVGLLFPPIQKSLTYC 480

Qy 481 FLVPTATICTGAIVLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540
 Db 481 FLVPTATICTGAIVLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540

Qy 541 TTPNTAWIQAAATTTATKKEHPL 563
 Db 541 TTPNTAWIQAAATTTATKKEHPL 563

RESULT 2
 AAB30521
 ID AAB30521 standard; protein; 563 AA.
 XX
 AC AAB30521;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE Amino acid sequence of a human GLUTX polypeptide.
 XX
 KW Human; GLUTX; glucose transporter; hexose transport;
 KW tissue-specific marker.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Domain 52..71
 FT /note= "transmembrane domain"
 FT Domain 108..128
 FT /note= "transmembrane domain"
 FT Domain 141..159
 FT /note= "transmembrane domain"
 FT Domain 166..189
 FT /note= "transmembrane domain"
 FT Domain 204..221
 FT /note= "transmembrane domain"
 FT Domain 233..252
 FT /note= "transmembrane domain"
 FT Domain 317..333
 FT /note= "transmembrane domain"
 FT Domain 355..375
 FT /note= "transmembrane domain"
 FT Domain 383..404
 FT /note= "transmembrane domain"
 FT Domain 413..437
 FT /note= "transmembrane domain"
 FT Domain 449..472
 FT /note= "transmembrane domain"
 FT Domain 481..499
 FT /note= "transmembrane domain"
 XX
 US6136547-A.
 PN
 XX 24-OCT-2000.
 PD
 XX
 XX 26-APR-1999; 99US-00299549.
 PF
 XX 26-FEB-1998; 98US-00031392.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 PA Weng X, Tartaglia LA;
 XX
 PI WPI; 2001-040236/05.
 DR N-PSDB; AAC62143.
 XX
 XX New human glucose transporter protein and nucleic acid encoding the
 PT protein, useful in the diagnosis and treatment of disorders associated
 PT with aberrant hexose transport.
 XX
 XX Claim 1; Fig 2A-D; 48pp; English.
 PS
 XX
 XX The present sequence represents a human GLUTX polypeptide. GLUTX is a
 CC glucose transporter protein. The GLUTX polypeptides and polynucleotides
 CC are useful in the diagnosis and treatment of disorders associated with
 CC aberrant hexose transport. The GLUTX polypeptide is useful for
 CC determining whether a given disorder is associated with aberrant
 CC expression of GLUTX or activity of GLUTX. The nucleic acid molecule and
 CC the GLUTX polypeptide are useful as diagnostic or therapeutic agents, or
 CC they can be used to generate antibodies or identify small molecules that,
 CC in turn are clinically useful. The GLUTX nucleic acid molecules are
 CC useful for identifying the chromosomal location of GLUTX and as tissue-

OM protein - protein search, using sw model

Run on: April 6, 2004, 09:59:52 ; Search time 59 Seconds
(without alignments)
2696.174 Million cell updates/sec

Title: US-09-981-947B-2
Perfect score: 2903
Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWIQAAATTTATKEHPL 563

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 25Jan04 :
1: geneseqp1980s : *
2: geneseqp1990s : *
3: geneseqp2000s : *
4: geneseqp2001s : *
5: geneseqp2002s : *
6: geneseqp2003as : *
7: geneseqp2003bs : *
8: geneseqp2004s : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2903	100.0	563	2	AAV27286 Human glu
2	2903	100.0	563	4	AB30521 Amino aci
3	2903	100.0	563	5	AAE14611 Human glu
4	2737	94.3	558	4	ABB11910 Human GLU
5	2737	94.3	558	4	AAW79422 Human pro
6	2720	94.0	558	4	AAW41116 Human pol
7	2723	93.8	537	5	AAE16788 Human tra
8	2527.5	87.1	507	7	ADC26251 Human NOV
9	2396	82.5	472	5	ABP52206 Human fac
10	2341	80.6	480	7	ADC26255 Human NOV
11	2337.5	80.5	565	7	ADC26253 Human NOV
12	2018	69.5	408	4	AAW78438 Human pro
13	2017	69.5	422	7	ADC26265 Human NOV
14	2012	69.3	408	4	AAW39530 Human pol
15	1939	66.8	416	4	AAE04893 Human tra
16	1929	66.4	396	7	ADC26261 Human NOV
17	1920	66.1	396	7	ADC26259 Human NOV
18	1911	65.8	396	7	ADC26263 Human NOV
19	1721	59.3	362	7	ADC26257 Human NOV
20	1122.5	38.7	501	5	ABP65173 Hypoxia-r
21	1122.5	38.7	501	6	ABP58597 Human can
22	1122.5	38.7	501	7	ABR56199 Human glu
23	1122.5	38.7	501	7	ADD01473 Human GLU
24	1119.5	38.6	512	5	AAO14130 Human tra
25	1119.5	38.6	512	7	ABR56198 Human glu

26	1119.5	38.6	512	7	ADD01365	Add01365 Human glu
27	1117	38.5	521	7	ADD01387	Add01387 Human var
28	1108	38.2	500	2	AAV27291	AAV27291 Glucose t
29	1081	37.2	493	4	AAE30522	AAE30522 Glucose t
30	1081	37.2	493	4	AAE30522	AAE30522 Amino aci
31	1073	37.0	565	6	AAE35305	AAE35305 Human 637
32	1001.5	34.5	500	5	ABR80602	ABR80602 Human sbg
33	1000.5	34.5	499	5	AAO14198	AAO14198 Human tra
34	999	34.4	496	4	AAE88405	AAE88405 Human mem
35	999	34.4	496	4	ABR80601	ABR80601 Human sbg
36	997.5	34.4	483	5	ABG70393	ABG70393 Human glu
37	979	33.7	524	6	ABG70513	ABG70513 Human 324
38	952.5	32.8	486	6	ABO00640	ABO00640 Novel hum
39	936.5	32.3	492	2	AAW17835	AAW17835 Human glu
40	936.5	32.3	492	5	ABP65129	ABP65129 Hypoxia-r
41	936.5	32.3	492	6	ABU56445	ABU56445 Lung canc
42	936	32.2	471	5	ABG70347	ABG70347 Novel hum
43	935	32.2	501	5	ABG70348	ABG70348 Novel hum
44	934.5	32.2	480	6	ABJ37062	ABJ37062 Human bre
45	934.5	32.2	492	2	AAE11360	AAE11360 Glucose T

ALIGNMENTS

RESULT 1
AAV27286
ID AAV27286 standard; protein; 563 AA.
XX
AC AAV27286;
XX
DT 05-NOV-1999 (first entry)
XX
DE Human glucose transporter protein GLUTX.
XX
KW Human; glucose transporter; GLUTX; hexose transport; gene therapy;
KW chromosome mapping.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 52..71
FT Domain /note= "transmembrane domain 1"
FT Domain 108..128
FT Domain /note= "transmembrane domain 2"
FT Domain 141..159
FT Domain /note= "transmembrane domain 3"
FT Domain 166..189
FT Domain /note= "transmembrane domain 4"
FT Domain 204..221
FT Domain /note= "transmembrane domain 5"
FT Domain 233..252
FT Domain /note= "transmembrane domain 6"
FT Domain 317..333
FT Domain /note= "transmembrane domain 7"
FT Domain 355..375
FT Domain /note= "transmembrane domain 8"
FT Domain 383..404
FT Domain /note= "transmembrane domain 9"
FT Domain 413..437
FT Domain /note= "transmembrane domain 10"
FT Domain 449..472
FT Domain /note= "transmembrane domain 11"
FT Domain 481..499
FT Domain /note= "transmembrane domain 12"
XX
US5942398-A.
XX
PD 24-AUG-1999.
XX
PF 26-FEB-1998;
XX
PR 26-FEB-1998;
98US-00031392.
98US-00031392.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Tartaglia LA, Weng X;

XX XX WPI; 1999-526192/44.

DR N-PSDB; AAZ11502.

XX A novel mammalian transporter protein and the gene encoding it, are

PT useful in the diagnosis and treatment of disorders associated with

PT aberrant sugar transport.

XX PS Claim 1; Fig 2A-D; 48pp; English.

XX This represents a human glucose transporter (GLUTX) protein. The protein

CC can be expressed by standard recombinant methodology. GLUTX nucleic acids

CC are useful as hybridization probes for detecting the presence of GLUTX

CC DNA in a sample, useful for diagnosing conditions associated with

CC aberrant expression levels of GLUTX. The GLUTX gene is also useful as a

CC therapeutic agent for regulating translation of GLUTX mRNA, and for

CC treatment of disorders associated with aberrant expression of GLUTX and

CC aberrant hexose transport. It is useful for generating GLUTX specific

CC antibodies, identifying agonists and antagonists of GLUTX, and

CC identifying nucleic acids in other species encoding nucleic acids

CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying

CC the chromosomal location of GLUTX, and as tissue specific markers

XX SQ Sequence 563 AA;

Query Match 100.0%; Score 2903; DB 2; Length 563;

Best Local Similarity 100.0%; Pred. No. 5.3e-287;

Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRRALLSCDHLRSVPGRRRKQWSCSLLVAS 60

DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRRALLSCDHLRSVPGRRRKQWSCSLLVAS 60

QY 61 LAGAGSSFLYGVNLSVNVNAPYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120

DB 61 LAGAGSSFLYGVNLSVNVNAPYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120

QY 121 GLVGTLLVVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLVGRFIMGDGV 180

DB 121 GLVGTLLVVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLVGRFIMGDGV 180

QY 181 ALSVLPWYLSEISPKIRSGSLQVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIW 240

DB 181 ALSVLPWYLSEISPKIRSGSLQVTAIFICIGVFTQGLLGLPELLGKSTWPLYFGVIW 240

QY 241 PAVVQLLSLFLPDSRPLYLLEKHNARAVKAFQTELGKADVSQVEEVLAEHVQSRIR 300

DB 241 PAVVQLLSLFLPDSRPLYLLEKHNARAVKAFQTELGKADVSQVEEVLAEHVQSRIR 300

QY 301 LVSVLELLRAPYVRWQVTVITMACYQCLGNALWFTYNSIFGKAGIPPAKIPYVTLST 360

DB 301 LVSVLELLRAPYVRWQVTVITMACYQCLGNALWFTYNSIFGKAGIPPAKIPYVTLST 360

QY 361 GGIETLAAPVSGLVIEHLGRRLPILGGFGLMGLFFGTITITLQDHAPWVPLSIVGIL 420

DB 361 GGIETLAAPVSGLVIEHLGRRLPILGGFGLMGLFFGTITITLQDHAPWVPLSIVGIL 420

QY 421 AIIASPCSGPGIPFILTGEFFQOSQRPAAFIAGTVNMLSNFPAVGLLFPFTQKSLDTC 480

DB 421 AIIASPCSGPGIPFILTGEFFQOSQRPAAFIAGTVNMLSNFPAVGLLFPFTQKSLDTC 480

QY 481 FLVPATICIGALYLVPLPETKNRTVAETISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540

DB 481 FLVPATICIGALYLVPLPETKNRTVAETISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540

QY 541 TTPNTAMIQAAATTTATKKEHPL 563

DB 541 TTPNTAMIQAAATTTATKKEHPL 563

RESULT 2

AAB30521

ID AAB30521 standard; protein; 563 AA.

XX AC AAB30521;

XX XX 06-MAR-2001 (first entry)

DT Amino acid sequence of a human GLUTX polypeptide.

XX Human; GLUTX; glucose transporter; hexose transport;

XX tissue-specific marker.

XX Homo sapiens.

OS Key Location/Qualifiers

XX 52..71

FT Domain /note= "transmembrane domain"

FT Domain 108..128

FT Domain /note= "transmembrane domain"

FT Domain 141..159

FT Domain /note= "transmembrane domain"

FT Domain 166..189

FT Domain /note= "transmembrane domain"

FT Domain 204..221

FT Domain /note= "transmembrane domain"

FT Domain 233..252

FT Domain /note= "transmembrane domain"

FT Domain 317..333

FT Domain /note= "transmembrane domain"

FT Domain 355..375

FT Domain /note= "transmembrane domain"

FT Domain 383..404

FT Domain /note= "transmembrane domain"

FT Domain 413..437

FT Domain /note= "transmembrane domain"

FT Domain 449..472

FT Domain /note= "transmembrane domain"

FT Domain 481..499

FT Domain /note= "transmembrane domain"

XX US136547-A.

FN 24-OCT-2000.

XX 26-APR-1999; 99US-00299549.

XX 26-FEB-1998; 98US-00031392.

PR (MILL-) MILLENNIUM PHARM INC.

XX Weng X, Tartaglia LA;

XX WPI; 2001-040236/05.

DR N-PSDB; AAC62143.

XX New human glucose transporter protein and nucleic acid encoding the

PT protein, useful in the diagnosis and treatment of disorders associated

PT with aberrant hexose transport.

XX Claim 1; Fig 2A-D; 48pp; English.

PS The present sequence represents a human GLUTX polypeptide. GLUTX is a

CC glucose transporter protein. The GLUTX polypeptides and polynucleotides

CC are useful in the diagnosis and treatment of disorders associated with

CC aberrant hexose transport. The GLUTX polypeptide is useful for

CC determining whether a given disorder is associated with aberrant

CC expression of GLUTX or activity of GLUTX. The nucleic acid molecule and

CC the GLUTX polypeptide are useful as diagnostic or therapeutic agents, or

CC they can be used to generate antibodies or identify small molecules that,

CC in turn are clinically useful. The GLUTX nucleic acid molecules are

CC useful for identifying the chromosomal location of GLUTX and as tissue-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 09:59:52 ; Search time 59 seconds

(without alignments)
2696.174 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903

Sequence: 1 MARKQNRNSELGLVPLTDD.....NTAWIQAAATTTATKKEHPL 563

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2903	100.0	563	2	AAY27286 Human glu
2	2903	100.0	563	4	AAB30521 Amino aci
3	2903	100.0	563	5	AAE14611 Human glu
4	2737	94.3	558	4	ABB11910 Human GLU
5	2737	94.3	558	4	AAM79422 Human pro
6	2730	94.0	558	4	AAM41316 Human pol
7	2723	93.8	537	5	AAE16788 Human tra
8	2597.5	87.1	507	7	ADC26251 Human NOV
9	2396	82.5	472	5	ABP52206 Human fac
10	2341	80.6	480	7	ADC26255 Human NOV
11	2337.5	80.5	565	7	ADC26253 Human NOV
12	2018	69.5	408	4	AAM78438 Human pro
13	2017	69.5	422	7	ADC26265 Human NOV
14	2012	69.3	408	4	AAM39530 Human tra
15	1939	66.8	416	4	AAE04893 Human tra
16	1929	66.4	396	7	ADC26261 Human NOV
17	1920	66.1	396	7	ADC26259 Human NOV
18	1911	65.8	396	7	ADC26263 Human NOV
19	1721	59.3	362	7	ADC26257 Human NOV
20	1122.5	38.7	501	6	ABP65173 Hypoxia-x
21	1122.5	38.7	501	6	ABR58597 Human can
22	1122.5	38.7	501	7	ABR56199 Human glu
23	1122.5	38.7	501	7	ADD01473 Human GLU
24	1119.5	38.6	512	5	AAO14190 Human tra
25	1119.5	38.6	512	7	ABR56198 Human glu

26	1119.5	38.6	512	7	ADD01365	Add01365 Human glu
27	1117	38.5	521	7	ADD01387	Add01387 Human var
28	1108	38.2	500	2	AAY27291	Aay27291 Glucose t
29	1081	37.2	493	2	AAY27292	Aay27292 Glucose t
30	1081	37.2	493	4	AAB30522	Aab30522 Amino aci
31	1073	37.0	565	6	AAE33305	Aae33305 Human 637
32	1001.5	34.5	500	5	ABB80602	Abb80602 Human sbg
33	1000.5	34.5	499	5	AAO14198	Aao14198 Human tra
34	999	34.4	496	4	AAE88405	Aae88405 Human mem
35	999	34.4	496	5	ABB80601	Abb80601 Human sbg
36	997.5	34.4	483	5	ABG70393	Abg70393 Human glu
37	979	33.7	524	6	ABG70513	Abg70513 Human 324
38	952.5	32.8	486	6	ABO00640	Ab000640 Novel hum
39	936.5	32.3	492	2	AAW17835	Aaw17835 Human glu
40	936.5	32.3	492	4	ABP65129	Abp65129 Hypoxia-x
41	936.5	32.3	492	6	ABU56445	Abu56445 Lung canc
42	936	32.2	471	5	ABG70347	Abg70347 Novel hum
43	935	32.2	501	5	ABG70348	Abg70348 Novel hum
44	934.5	32.2	480	6	ABJ37062	Abj37062 Human bre
45	934.5	32.2	492	2	AAR11360	Aar11360 Glucose T

ALIGNMENTS

RESULT 1
AAY27286

ID AAY27286 standard; protein; 563 AA.

XX

XX AAY27286;

DT 05-NOV-1999 (first entry)

XX

DE Human glucose transporter protein GLUTX.

XX

KW Human; glucose transporter; GLUTX; hexose transport; gene therapy; chromosome mapping.

XX

OS Homo sapiens.

XX

EH Key Location/Qualifiers

FT Domain 52..71

FT Domain /note= "transmembrane domain 1"

FT Domain 108..128

FT Domain /note= "transmembrane domain 2"

FT Domain 141..159

FT Domain /note= "transmembrane domain 3"

FT Domain 166..189

FT Domain /note= "transmembrane domain 4"

FT Domain 204..221

FT Domain /note= "transmembrane domain 5"

FT Domain 233..252

FT Domain /note= "transmembrane domain 6"

FT Domain 317..333

FT Domain /note= "transmembrane domain 7"

FT Domain 355..375

FT Domain /note= "transmembrane domain 8"

FT Domain 383..404

FT Domain /note= "transmembrane domain 9"

FT Domain 413..437

FT Domain /note= "transmembrane domain 10"

FT Domain 449..472

FT Domain /note= "transmembrane domain 11"

FT Domain 481..499

FT Domain /note= "transmembrane domain 12"

XX US5942398-A.

PN 24-AUG-1999.

XX 98US-00031392.

PD 26-FEB-1998;

XX 98US-00031392.

PR 26-FEB-1998;

327	CCCCGTACATCAAGGCCCTTTTACAATGAGTTCATGGGAAAGGAAGGCATGGACATGCCAATAG	386
Db		
377	ACCCAGACACTCTGACTCTGCTCTGGTCTGTGCACCTGTGTCCATATTGCGCATCGGTGGAC	436
Qy		
387	ACCCAGACACTCTGACTCTGCTCTGGTCTGTGCACCTGTGTCCATATTGCGCATCGGTGGAC	446
Db		
437	TTGTGGGACGTTAA-TTGTGAAGATGATTGAAAGGTTCTTTGGGAGG--AAGCACACT	492
Qy		
447	TTGTGGGACATTAACTTTGTGAAGATGATTGAAAGGTTCTTTGGGAGGAAACACACCTT	506
Db		
493	TTGTCTGGCCAAATAATGGGTTTGCATTT	520
Qy		
507	TGGTGGGGCCATTAAATGGGGTTTGCATTT	534
Db		

RESULT 15
 BE745805
 LOCUS
 601580005F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928801 5',
 DEFINITION
 mRNA sequence.
 ACCESSION
 BE745805
 VERSION
 BE745805.1 GI:10159797
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation by: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
plate: LLCW760 row: j column: 02
High quality sequence stop: 752.

FEATURES	Location/Qualifiers
source	1. .961

```

/db_xref="taxon:9606"
/clone="IMAGE:3928801"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site 1: xhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)".

```

ORIGIN

Query Match 18.8%; Score 439.4; DB 10; Length 961;
Best Local Similarity 97.2%; Pred. No. 4e-51;
Matches 479; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY	217	AGAAAGACTGTCCTGCTCGTCTCGTGGCTCCCTCGCGGGCGCTTGGCTCTCTCC	276
n ^b	274	AATTGGACTGTGCTCGTCTCGTGGCTCCCTCGCGGGCGCTTGGCTCTCTCC	333

33

QY
277 TTCTCTACGGCTACAACCTGTCTGGTGGTGAATGCCCCACCCCGTACATCAAGGCC¹¹¹ 336

393

DB
334 TTCCCTACGGCTACAACCTGTCGGTGGTGATGCCCAAGCCCCCGA

337 TACAATGAGTCATGGGAAGAAGGCATGGACGTCCTCAATAGACCCAGACACTCTGACTCTG 396

Query Match 19.3%; Score 452.8; DB 12; Length 1297;
Best Local Similarity 95.9%; Pred. No. 5.1e-53;
Matches 487; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

17 GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGTCACTGAGACCCATGG 76

100

Db 27 GCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 86

136

QY CAAGGAAACAAAAATAGGAATTCATAGGATATCGGGCCATAGTTCCTCCACAGATGCGACGCTTC

87 CAAGGAAACAAAATAGGAATTCCAAGGAAGCTGGGCCTAGTTCCCCTCACAGATGACACCA 146

100

QY 137 GCCACGCCGGCCTCCAGGCCAGGGCACTGCTGGAGTGTGACCACTGAGGAGTG 196

CCCTGACCGTGGAGGATG

DB T47 GGCACGCCAGGCCCTCCAGGGCCACGGGAGGGGACACGCTGCGAGAGTGTGACCCCTGAGGCGATG

197 GGGTGCCAGGTGGAAAGGAGAAGGACTGGTCTGTCTCGTCTCGTGGCTCCCTCG 256

D_b 207 GGGTGCCAGGTGAAGGAGAGAAAGGACTGGTCTGCTGCTCCTCGTGGCCTCCCTCG 266

[illegible]

QY
25 / CCGGCGCCCTTCGGGTCCCTCCTTCCCTACAGGCACACACCTGTCGGCTGGATTCGCCCA

267 CGGGCGCCTTCGGCTCCTCCTTCTACGGCTACAACCTGTCGGTGGTGAATGCCCCCA 326

100

317 CCCGTCATCAAGGCCTTTTACAATGAGTCATGGGAAAGAGGCATGGACGTCCTCAATAG 376

Matches 587; Conservative 0; Mismatches 112; Indels 1; Gaps 1;
Qy 950 GCACAGTGCAGAGAGATCCGCTGTGTCTGCTGTGAGCTGTGAGAGTCCCTACG 1009
Db 1 GCGCGTGCAGAGAGATCCGCTGTGTCTGCTGTGAGCTGTGAGAGTCCCTACG 60
Qy 1010 TCCGCTGGCAGT 1069
Db 61 TCCGCTGGCAGT 120
Qy 1070 ATGCAATTTGGT 1129
Db 121 ATGGGATCTGGT 180
Qy 1130 TCCCATACGTCACTTGT 1189
Db 181 TCCCATACGTCACTTGT 240
Qy 1190 TGGTCAATGAGCCTGGGAGGAGAGACCTCTCTCATGTGTGTGTGTGTGTGTGTGT 1249
Db 241 TGGTCAATGAGCCTGGGAGGAGAGACCTCTCTCATGTGTGTGTGTGTGTGTGTGT 300
Qy 1250 TCTTCTTTGGACCTTCAACATCAGCTGACCTGAGGACCAAGCCCTGCGTCCCTT 1309
Db 301 TCTTCTTTGGACCTTCAACATCAGCTGACCTGAGGACCAAGCTTCCCTGCGTCCCTT 360
Qy 1310 ACCTGAGT 1369
Db 361 ACCTGAGT 420
Qy 1370 TCCCGTTTCATTTGAGT 1429
Db 421 TCCCATTCCTTGACCGGAGGTTCTTCAACAGTCAAGAGCAGCGGAGCGCTTCATGA 480
Qy 1430 TTGAGGACCGTCAACTGT 1489
Db 481 TCGCAGGACAGTCAACTGT 540
Qy 1490 AGAAAGTGTGAGACCTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1549
Db 541 AGAAAGTGTGAGACCTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 1550 TCTACCTGTATTTTGT 1609
Db 601 CCTACTTCTATTTGTCTTCCCGGAGACCAAGACAGAACCC-ATGCAGAAATCAGCCAGG 659
Qy 1610 CATTTTCCAAAGGAAACAAAGCATATCCACAGAGAGAA 1649
Db 660 CCATTGCCAGAGGAAACAAAGCTCAGACCCAGAGGTGAA 699

RESULT 13
BE740763
LOCUS 601593559F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3947511 5',
DEFINITION BE740763 893 bp mRNA linear EST 15-SEP-2000
mRNA sequence.
ACCESSION BE740763
VERSION BE740763.1 GI:10154755
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 893)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

FEATURES
source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM809 row: e column: 16
High quality sequence stop: 734.
Location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3947511"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 20.1%; Score 471.6; DB 10; Length 893;
Best Local Similarity 90.6%; Pred. No. 1.5e-55;
Matches 560; Conservative 0; Mismatches 49; Indels 9; Gaps 5;
Qy 217 AGAAAGGAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 276
Db 202 AAATTGGACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 261
Qy 277 TTCTCTTACGGTACAACCTGT 336
Db 262 TTCTCTTACGGTACAACCTGT 321
Qy 337 TACATGAGTCAATGGGAAGAAGGATGAGCGTCCCAATAGACCCAGACACTCTGACTCTG 396
Db 322 TACATGAGTCAATGGGAAGAAGGATGAGCGTCCCAATAGACCCAGACACTCTGACTCTG 381
Qy 397 CTCTGTCTGTGACTGTGTCTATATTCGCCATTCGGTGGACTTGTGGGACGTTAATTGTG 456
Db 382 CTCTGTCTGTGACTGTGTCTATATTCGCCATTCGGTGGACTTGTGGGACGTTAATTGTG 441
Qy 457 AAGATGATTGGAAGGTTCTTGGGAGGAGACACATTTGTGGGCCAAATAATGGGTTTGA 516
Db 442 AAGATGATTGGAAGGTTCTTGGGAGGAGACACATTTGTGGGCCAAATAATGGGTTTGA 501
Qy 517 ATTTCTGTGCAATGCTGATGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 576
Db 502 ATTTCTGTGCAATGCTGATGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 561
Qy 577 GTGGACCGTTTCATCATGCGCATAGATGAGGCGCTCGCCTCAGTGTGTCTGCCCATGTAC 636
Db 562 GTGGACCGTTTCATCATGCGCATAGATGAGGCGCTCGCCTCAGTGTGTCTGCCCATGTAC 620
Qy 637 CTCAGTGAGATCTCACCCCAAGG--AGATCCGTGGTCTCTGTGGGCGAGGTGACTG-CCATC 693
Db 621 CTCAGTGAGATCTCACCCCAAGG--AGATCCGTGGTCTCTGTGGGCGAGGTGACTGCCCCATC 680
Qy 694 TTTATCTGCAATGGCGTGTTCAC---TGGGACGTTCTGTGGGCTGTGGGCTGTGGGCTGTGGG 750
Db 681 TTTATCTGCAATGGCGTGTTCACGTTGGGCGAGTTCTTGGGCTGTGGGCTGTGGGCTGTGGG 740
Qy 751 AAGGAGAGTACTGCGCATACCTGTTTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 810
Db 741 GACAGGGAGTCTCTGCGCATACCTGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 798
Qy 811 CTGAGCGCTTCCCTTCTC 828
Db 799 GTGGGCGCTTCCCTTCTC 816

RESULT 14

654 G-CAGGTGGCATCCCGTTCACTGGACTGGTGGTAGTCTTCCACGCACTCTCAGCGGG 711

RESULT 11
 BE741795 1017 bp mRNA linear EST 15-SEP-2000
 LOCUS 60159563F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949549 5',
 DEFINITION mRNA sequence.

ACCESSION BE741795
 VERSION BE741795.1 GI:10155787
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL 1 (Bases 1 to 1017)
 COMMENT NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LiCM814 row: j column: 14
 High quality sequence start: 34
 High quality sequence stop: 828.
 Location/Qualifiers
 1..1017
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3949549"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 9"
 /note="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 24.9%; Score 584; DB 10; Length 1017;
 Best Local Similarity 98.1%; Pred. No. 5.2e-71;
 Matches 612; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 217 AGAAAGGACTGGTCTGCTCTGCTCTCTCGTGGCTTCCTCTCGGGGGCCCTTGGCTCTCTCC 276
 Db 252 AAATTGGACTGGTCTGCTCTGCTCTCTCGTGGCTTCCTCTCGGGGGGGCCCTTGGCTCTCTCC 311

QY 277 TTCCTCTAGGGGTACAACTGTCGGTGGTGAATGCCCGCCACCCCGTACATCAAGGCCCTTT 336
 Db 312 TTTCCTCTACGGGTACAACCTGTCGGTGGTGAATGCCCGCCACCCCGTACATCAAGGCCCTTT 371

QY 337 TACAATGAGTCATGGGAAGGACGATGGACGTCCTCAATAGACCCAGACACTCTGACTCTG 396
 Db 372 TACAATGAGTCATGGGAAGGACGATGGACGTCCTCAATAGACCCAGACACTCTGACTCTG 431

QY 397 CTCCTGGTCTGTGACTGTGTCATATTCGCCATCGGTGGACTTGTGGGGACGTTAAATTGTG 456
 Db 432 CTCCTGGTCTGTGACTGTGTCATATTCGCCATCGGTGGACTTGTGGGGACATTAATTGTG 491

QY 457 AAGATGATGGAAAGGTTCTTGGGAGGAGCAGCACTTGTGTCGCCCAATTAATCGGTTTGA 516
 Db 492 AAGATGATGGAAAGGTTCTTGGGAGGAGCAGCACTTGTGTCGCCCAATTAATCGGTTTGA 551

REFERENCE
1 (bases 1 to 889)

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12929891.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2476.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI026CD12NP1&cluster=2476.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DI026CD12NP1.
FEATURES
source 1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI026YH23"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 33.0%; Score 773.4; DB 9; Length 940;
Best Local Similarity 99.6%; Pred. No. 5.1e-97;
Matches 774; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 898 TTCTTGGTAAAGCAGACAGCTTTCCCAAGAGGTAGAGAGTCTCTGCTGAGACCAAGTGG 957
DB 940 TTCTTGGTAAAGCAGACAGCTTTCCCAAGAGGTAGAGAGTCTCTGCTGAGACCAAGTGG 881
QY 958 CAGAGAGCATCCGCTGTGTCGCTGGTGGAGTCTCTGAGAGTCCCTACGTCGCTGG 1017
DB 880 CAGAGAGCATCCGCTGTGTCGCTGGTGGAGTCTCTGAGAGTCCCTACGTCGCTGG 821
QY 1018 CAGGTGTCACGTGATTTCTCACCATGGCTGCTACCATGCTGTGGCTCTCAATGCAATT 1077
DB 820 CAGGTGTCACGTGATTTCTCACCATGGCTGCTACCATGCTGTGGCTCTCAATGCAATT 761
QY 1078 TGCTTCTATACCAACAGCATCTTTGGAAGAGCTGGGATCCCTCCGGCAAGATCCCATAC 1137
DB 760 TGCTTCTATACCAACAGCATCTTTGGAAGAGCTGGGATCCCTCCGGCAAGATCCCATAC 701
QY 1138 GTCACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCGCTCTCTCTGTTGGTCAATT 1197
DB 700 GTCACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCGCTCTCTCTGTTGGTCAATT 641
QY 1198 GAGCAGCTGGGACGAGACCCCTCTCATGTGGGCTTTGGGCTCATGGGCTCTCTTT 1257
DB 640 GAGCAGCTGGGACGAGACCCCTCTCATGTGGGCTTTGGGCTCATGGGCTCTCTTT 581
QY 1258 GGAGCCTCACCATCAGCTGACCTGACGAGCAGCCGCTGGTCCCTACCTGAGT 1317
DB 580 GGAGCCTCACCATCAGCTGACCTGACGAGCAGCCGCTGGTCCCTACCTGAGT 521
QY 1318 ATCGTGGGATTTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTC 1377
DB 520 ATCGTGGGATTTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTC 461
QY 1378 ATCTTGAGTGGTGGTCTTTCCAGCAATCTCAGCGGCGGCTGCTTCATCATATTGAGGC 1437

Db 460 ATCTTGACTGGTGGTCTTCCAGCAATCTCAGGGCGGGCTGCTTCATCATTCATTCGAGGC 401
QY 1438 ACCGTCAACTGGTCTCCAACTTTGCTGTGGGCTCTCTTCCCATTCATTCAGAAAGT 1497
Db 400 ACCGTCAACTGGTCTCCAACTTTGCTGTGGGCTCTCTTCCCATTCATTCAGAAAGT 341
QY 1498 CTGAGACCACTACTGTTTCTTCCAGTCTTTGCTACAAATTTGTATCAGAGTGTATCTACTG 1557
Db 340 CTGAGACCACTACTGTTTCTTCCAGTCTTTGCTACAAATTTGTATCAGAGTGTATCTACTG 281
QY 1558 TATTTTGTGCTGCTGAGACCAAAACAGAAACCTTATCGAGAAATCAGCAGGCAATTTTCC 1617
Db 280 TATTTTGTGCTGCTGAGACCAAAACAGAAACCTTATCGAGAAATCAGCAGGCAATTTTCC 221
QY 1618 AAAAGGAACAAAGCATATCCACCAGAGAGAGAAATCGACTCAGCTGTCTACTGTATGCT 1674
Db 220 AAAAGGAACAAAGCATATCCACCAGAGAGAGAAATCGACTCAGCTGTCTACTGTATGCT 164
RESULT 6
LOCUS BI760988
DEFINITION 603048508P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188849 5', mRNA sequence.
ACCESSION BI760988
VERSION BI760988.1 GI:15752566
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbe@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L14M1471 row: p column: 02
High quality sequence start: 5
High quality sequence stop: 859.
FEATURES
source 1..876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5188849"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/notes="Organ: Pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 32.3%; Score 757.2; DB 12; Length 876;
Best Local Similarity 96.0%; Pred. No. 8.8e-95;
Matches 831; Conservative 0; Mismatches 28; Indels 7; Gaps 5;
QY 79 AGGAACAAATAGGAATTCCAA-GGAAGTGGGCTAGT-TCCCTCAGATGACACCA 136

RESULT 5
AL572022/c
LOCUS
AL572022
DEFINITION
AL572022 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODIO26YH23 3-PRIME, mRNA sequence.
ACCESSION
AL572022
VERSION
AL572022.2 GI:31293412
KEYWORDS
EST.

FEATURES

source

1. 932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1026VH23"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 36.4%; Score 852.2; DB 9; Length 932;
Best Local Similarity 99.6%; Pred. No. 7.6e-108;
Matches 854; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 17 GCCTGGCAGAGTCTGGGCTCCCTGGACTAGCCATCAGCTGGTCACTGAGACCCATGG 76
DB 76 GCCTGGCAGAGTCTGGGCTCCCTGGACTAGCCATCAGCTGGTCACTGAGACCCATGG 135
QY 77 CAAGGAAACAAATAGGAATCCAGGAATCTGGGCTAGTTCCTCTACAGATCACACCA 136
DB 136 CAAGGAAACAAATAGGAATCCAGGAATCTGGGCTAGTTCCTCTACAGATCACACCA 195
QY 137 GCCACGCCGGCTCCAGGGCCAGGGAGGACACTGCTGAGTGTGACCACTGAGGAGTG 196
DB 196 GCCACGCCGGCTCCAGGGCCAGGGAGGACACTGCTGAGTGTGACCACTGAGGAGTG 255
QY 197 GGGTCCAGGTGGAAGAGAGAAAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
DB 256 GGGTCCAGGTGGAAGAGAGAAAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 315
QY 257 GGGGGCCCTGGGCTCCCTCTCTACGCTCAACCTGCTGCTGCTGCTGCTGCTGCTG 316
DB 316 GGGGGCCCTGGGCTCCCTCTCTACGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCT 375
QY 317 CCCCTACATCAAGCCCTTTACATGATGATGATGGAAGAGGATGAGCTGCAATAG 376
DB 376 CCCCTACATCAAGCCCTTTACATGATGATGATGGAAGAGGATGAGCTGCAATAG 435
QY 377 ACCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB 436 ACCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
QY 437 TTGTGGGACGTTAATCTGAGATGATGAAAGGTTCTTGGAGGAGACACTTTGC 496
DB 496 TTGTGGGACATTAATCTGAGATGATGAAAGGTTCTTGGAGGAGACACTTTGC 555
QY 497 TGGCAATAATGGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
DB 556 TGGCAATAATGGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
QY 557 GAGCCTTTGAAATGCTCATTGTGGGACGCTTCATCATGGGATAGATGGGCGTGCCTC 616
DB 616 GAGCCTTTGAAATGCTCATTGTGGGACGCTTCATCATGGGATAGATGGGCGTGCCTC 675
QY 617 TCAGTGTCTCCCATGCTACCTAGTGAATCTCACCAAGGATCCGTGGCTCTCTGG 676
DB 676 TCAGTGTCTCCCATGCTACCTAGTGAATCTCACCAAGGATCCGTGGCTCTCTGG 735
QY 677 GGCAGGTGACCTGCTTATCTGCAATGCGGTGTTCACTGGGACGCTTCTGGGCTGTC 736
DB 736 GGCAGGTGACCTGCTTATCTGCAATGCGGTGTTCACTGGGACGCTTCTGGGCTGTC 795
QY 737 CCGAGCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATGCTGCTGCT 796
DB 796 CCGAGCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATGCTGCTGCT 855
QY 797 CCGTGTCCAGCTGTGAGCCTTCCCTTTCTCCGGACAGCCACGCTACCTGCTCTTGG 856

Db

QY

Db

RESULT 2

BM545063

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1131)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12360 row: k column: 23

High quality sequence start: 28

High quality sequence stop: 725.

FEATURES

source

1. 1131

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:558878"

/lab_host="DH10B"

/clone_lib="NIH_MGC_125"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;

Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool

of three ovaries, from females ranging in age from 38 to

49 yo. Library is oligo-dT primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 2.1 kb, insert size range 1-3.5 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 036."

ORIGIN

Query Match 36.2%; Score 848.8; DB 12; Length 1131;

Best Local Similarity 96.0%; Pred. No. 2e-107;

Matches 968; Conservative 0; Mismatches 27; Indels 13; Gaps 9;

QY

Db

QY

Db

QY

Db

QY

Db

QY

856 CCCTTGTCCAGTCTGAGCCTTCCCTTCTCCGACAGCCACGCTACCTGCTTGG 915
857 AGAAGCACACGAGGCA 873
916 AGAAGCACACGAGGCA 932

BM545063 1131 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6497736 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:558878
5', mRNA sequence.

BM545063
BM545063.1 GI:18776860
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1131)
NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12360 row: k column: 23
High quality sequence start: 28
High quality sequence stop: 725.

Location/Qualifiers
1. 1131
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:558878"
/lab_host="DH10B"
/clone_lib="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

Query Match 36.2%; Score 848.8; DB 12; Length 1131;
Best Local Similarity 96.0%; Pred. No. 2e-107;
Matches 968; Conservative 0; Mismatches 27; Indels 13; Gaps 9;

313 CCCACCCCTACATCAAGCCCTTTACATGATGATGGAAGAGGATGAGCTGCA 372
28 CCCACCCCTACTTC-AGGCCCTTTACATGATGATGGAAGAGGATGAGCTGCA 86
373 ATAGACCCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
87 ATAGACCCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 146
433 GGACTTGTGGGACCGTTAATTGTAAGATGATTGGAAGGTTCTTGGGAGGAGCACT 492
147 GGACTTGTGGGACCGTTAATTGTAAGATGATTGGAAGGTTCTTGGGAGGAGCACT 206
493 TTGCTGGCCAAATATGGGTTTCAATTTCTGCTGATGCTGATGCTGCTGCTGCTG 552

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 22:38:47 ; Search time 5970 Seconds
(without alignments)
11719.780 Million cell updates/sec

Title: US-09-981-947B-1
Perfect score: 2343
Sequence: 1 tcgaccacgctcgccct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	852.2	36.4	932	9	AL546827
2	848.8	36.2	1131	12	BM545063
3	832	35.5	968	12	BI517809
4	778.2	33.2	929	12	BI911320

C	5	773.4	33.0	940	9	AL572022
	6	757.2	32.3	876	12	BI760988
	7	754.2	32.2	889	12	BI821650
	8	752.8	32.1	2901	11	AK037093
C	9	669.2	28.6	687	14	CA309954
	10	588.6	25.1	916	10	BF035577
	11	584	24.9	1017	10	BE741795
	12	508.8	21.7	801	12	BI154904
	13	471.6	20.1	893	10	BE740763
	14	452.8	19.3	1297	12	BM925071
	15	439.4	18.8	961	10	BE745805
	16	432.2	18.4	2473	11	AK033725
C	17	428.4	18.3	675	14	CA307279
	18	427.6	18.3	1539	29	AY401415
	19	424.4	18.1	3097	11	AK029720
C	20	422.4	18.0	690	12	BM678584
	21	419.6	17.9	2100	11	AF289610
	22	415.2	17.7	667	10	BE625357
	23	411.2	17.6	597	10	BE305606
	24	408.6	17.4	682	12	BG971798
C	25	408.4	17.4	620	12	BQ015082
C	26	407.4	17.4	595	10	AW300078
	27	403.6	17.2	602	9	AI766064
C	28	397.4	17.0	629	10	BB660933
C	29	391.6	16.7	605	10	BE463553
C	30	387.8	16.6	597	10	AW263084
	31	387.8	16.6	1542	29	AY401417
	32	378.8	16.2	635	10	BB660981
C	33	376.8	16.1	962	12	BI762013
C	34	375.4	16.0	571	9	AI671682
	35	368.8	15.7	641	10	BB664247
C	36	365.2	15.6	608	10	BE464196
C	37	356.2	15.2	563	9	AI433678
C	38	355.2	15.2	1506	29	AY410442
C	39	316.4	13.5	495	9	AI312682
C	40	313.2	13.4	602	10	BB630190
	41	311.6	13.3	770	14	CB521349
	42	306.6	13.1	339	14	FO0548
C	43	304.4	13.0	498	10	AW614815
C	44	302.8	12.9	500	9	AI022719
	45	300.4	12.8	1453	29	AY401416

ALIGNMENTS

RESULT 1
AL546827
LOCUS
DEFINITION AL546827 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI026VH23 5-PRIME, mRNA sequence.
ACCESSION AL546827
VERSION AL546827.2 GI:31268660
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 932)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12880321.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2476.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CSODI026CDI2QP1&cluster=2476.r. Contact : Feng Liang Email : fliang@lifetech.com URL :

AL572022 AL572022
BI760988 603048508
BI821650 603036688
AK037093 Mus muscu
CA309954 UI-H-FTL-
BF035577 601457471
BE741795 601595653
BI154904 602902896
BE740763 601593559
BM925071 AGENCOURT
BE745805 601580005
AK033725 Mus muscu
CA307279 UI-H-FTL-
AY401415 Homo sapi
AK029720 Mus muscu
BM678584 UI-E-EOO-
AF289610 Homo sapi
BE625357 BB625357
BE305606 601102363
BG971798 602840867
BQ015082 UI-H-ED1-
AW300078 xs45H04.x
AI766064 wh67e10.x
BB660933 BB660933
BE463553 hw25a06.x
AW263084 xm35f10.x
AY401417 Mus muscu
BB660981 BB660981
BI762013 603048963
AI671682 ty26f09.x
BB664247 BB664247
BE464196 hw85a05.x
AI433678 ti88b12.x
AY410442 Homo sapi
AI312682 qp83a08.x
BB630190 BB630190
CE521349 UI-M-GHO-
FO0548 HSB10E082 S
AW614815 h980a04.x
AI022719 ow31b01.x
AY401416 Pan trogl

Search completed: April 8, 2004, 22:43:36
Job time : 839 secs

QY 1261 ACCCTACCAATCAGCTGACCCCTGCAGGACCAAGCCCGCCCTGGTCCCTACCTGAGTATC 1320
Db 812 ACCCTACCAATCAGCTGACCCCTGCAGGACCAAGCCCGCCCTGGTCCCTACCTGAGTATC 871
QY 1321 GTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCCAGGTGGCATCCCGTTTATC 1380
Db 872 GTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCCAGGTGGCATCCCGTTTATC 931
QY 1381 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCCGGCTTGCCTTCATCATTTGCAAGCACC 1440
Db 932 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCCGGCTTGCCTTCATCATTTGCAAGCACC 991
QY 1441 GTCACTGGTCTCCAACTTTGTGTGGTGGTCTCTTCCCAATTCATTCAGAAAAAGTCTG 1500
Db 992 GTCACTGGTCTCCAACTTTGTGTGGTGGTCTCTTCCCAATTCATTCAGAAAAAGTCTG 1051
QY 1501 GACACCTACTGTTTCCCTAGCTTTTGTGTACAAATTTGTATACAGGTGGTCTTACCTGTAT 1560
Db 1052 GACACCTACTGTTTCCCTAGCTTTTGTGTACAAATTTGTATACAGGTGGTCTTACCTGTAT 1111
QY 1561 TTTGTGTCTGCTCAGACCAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCAAA 1620
Db 1112 TTTGTGTCTGCTCAGACCAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCGAG 1171
QY 1621 AGGAA 1625
Db 1172 GGCAA 1176

RESULT 13

US-10-188-186-87
; Sequence 87, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 87
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1189)
US-10-188-186-87

Query Match 48.3%; Score 1130.6; DB 12; Length 1189;

Best Local Similarity 99.2%; Pred. No. 9.2e-301;
Matches 1136; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 481 AGGAAGCAGACATTTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGCAATGCTGATGACC 540
Db 32 AGGAAGCAGACATTTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGCAATGCTGATGACC 91
QY 541 TGCTCTCTCCAGCAGGAGCCCTTTGAAATGCTCATTTGTGGACGCTTTCATATGGGCATA 600
Db 92 TGCTCTCTCCAGCAGGAGCCCTTTGAAATGCTCATTTGTGGACGCTTTCATATGGGCATA 151
QY 601 GATGAGGCGTTCGCTCAGTGTCTGCCCAATGTAACCTGATGAGATCTCAACCAAGAG 660
Db 152 GATGAGGCGTTCGCTCAGTGTCTGCCCAATGTAACCTGATGAGATCTCAACCAAGAG 211
QY 661 ATCCGTTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCATTTGGGCTGTTCACTGG 720
Db 212 ATCCGTTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCATTTGGGCTGTTCACTGG 271
QY 721 CAGCTTCTGGGCTCTGCCAGTCTGTGGGAAAGAGAGTACTGTGCCCATACCTGTTTGA 780
Db 272 CAGCTTCTGGGCTCTGCCAGTCTGTGGGAAAGAGAGTACTGTGCCCATACCTGTTTGA 331
QY 781 GTGATTGTGGTCCCTGGCGGTTCTCAGTGTGTGAGCTTCCCTTTCTCCGAGACAGGCCA 840
Db 332 GTGATTGTGGTCCCTGGCGGTTCTCAGTGTGTGAGCTTCCCTTTCTCCGAGACAGGCCA 391
QY 841 CGCTACCTGCTCTTGGAGAAACACAAACGAGGACAGAGCTGTGAAGCCCTTCCAAACGTT 900
Db 392 CGCTACCTGCTCTTGGAGAAACACAAACGAGGACAGAGCTGTGAAGCCCTTCCAAACGTT 451
QY 901 TTGGTAAAGCAGACGTTTCCACAGGTAGAGGAGTCTCTGGGTGAGAGCCACGTCGAG 960
Db 452 TTGGTAAAGCAGACGTTTCCACAGGTAGAGGAGTCTCTGGGTGAGAGCCACGTCGAG 511
QY 961 AGGAGCATCCGCTGGTGTCTGCTGAGTGTCTGAGAGCTCCCTAGCTGGTGGCAG 1020
Db 512 AGGAGCATCCGCTGGTGTCTGCTGAGTGTCTGAGAGCTCCCTAGCTGGTGGCAG 571
QY 1021 GTGGTCAACCGTCAATTTGACCAATGGCTGTACAGCTCTGTGGCTCAATGCAATTGG 1080
Db 572 GTGGTCAACCGTCAATTTGACCAATGGCTGTGTACAGCTCTGTGGCTCAATGCAATTGG 631
QY 1081 TTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCATAGTC 1140
Db 632 TTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCATAGTC 691
QY 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTGTTGGTGGTCAATTGAG 1200
Db 692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTGTTGGTGGTCAATTGAG 751
QY 1201 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 1260
Db 752 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 811
QY 1261 ACCCTCACCATCAGCTGACCTTGACGACCAACGCCCCCTGGGTCCCTACCTGAGTATC 1320
Db 812 ACCCTCACCATCAGCTGACCTTGACGACCAACGCCCCCTGGGTCCCTACCTGAGTATC 871
QY 1321 GTGGGATTTGGCCATCATGCTCTTTCTGAGTGGGCCAGGTGGCATCCCTTCATC 1380
Db 872 GTGGGATTTGGCCATCATGCTCTTTCTGAGTGGGCCAGGTGGCATCCCTTCATC 931
QY 1381 TTGACTGTGTAGTTCTTCCAGCAATCTCAGCGCGCGGCTGCTTTCATCATTCAGAGCACC 1440
Db 932 TTGACTGTGTAGTTCTTCCAGCAATCTCAGCGCGCGGCTGCTTTCATCATTCAGAGCACC 991
QY 1441 GTCAACTGGCTCTCCAACTTTGTGTGGTGGTCTCTTCCCAATTCATTCAGAAAAAGTCTG 1500
Db 992 GTCAACTGGCTCTCCAACTTTGTGTGGTGGTCTCTTCCCAATTCATTCAGAAAAAGTCTG 1051
QY 1501 GACACCTACTGTTTCTTGTGTACAAATTTGTATACAGGTGGTCTTACCTGTAT 1560

961 AGGAGCATCGGCTGCTGTCGCTGGAGCTGCTGAGAGCTCCCTAGCTCGCTGGCAG 1020
512 AGGAGCATCGGCTGCTGTCGCTGGAGCTGCTGAGAGCTCCCTAGCTCGCTGGCAG 571
1021 GTGGTCACCGTGAATGTACCAATGGCTGCTACCAAGCTCTGTGGGCTCAATGCAATTTGG 1080
572 GTGGTCACCGTGAATGTACCAATGGCTGCTACCAAGCTCTGTGGGCTCAATGCAATTTGG 631
1081 TTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 1140
632 TTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 691
1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTCTGTTTGGTCAATTGAG 1200
692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTCTGTTTGGTCAATTGAG 751
1201 CACCTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260
752 CACCTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 811
1261 ACCTCACCATCAGCTGACCTGACGAGACCAAGCCCTGCTGGTCCCTTACCTGAGTATC 1320
812 ACCTCACCATCAGCTGACCTGACGAGACCAAGCCCTGCTGGTCCCTTACCTGAGTATC 871
1321 GTGGGATCTCTGGCCATCATGCCCTCTTCTGAGTGGCCAGGTGGCATCCCGTTTCATC 1380
872 GTGGGATCTCTGGCCATCATGCCCTCTTCTGAGTGGCCAGGTGGCATCCCGTTTCATC 931
1381 TTGACTGTGTGAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCCATCATTCAGGACCC 1440
932 TTGACTGTGTGAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCCATCATTCAGGACCC 991
1441 GTCAACTGGCTCTCCAACTTTCTGTGGGCTCTCTTCCCATCATTCAGGAAAGTCTG 1500
992 GTCAACTGGCTCTCCAACTTTCTGTGGGCTCTCTTCCCATCATTCAGGAAAGTCTG 1051
1501 GACACCTACTGTTTCTAGTCTTTGTGTACAAATTTGTATACAGGTGCTATCTACCTGAT 1560
1052 GACACCTACTGTTTCTAGTCTTTGTGTACAAATTTGTATACAGGTGCTATCTACCTGAT 1111
1561 TTGTGCTCCCTGAGACCAAAACAGAACTATGAGAAATCAGACGAGCAATTTTCCAAA 1620
1112 TTGTGCTCCCTGAGACCAAAACAGAACTATGAGAAATCAGACGAGCAATTTTCCGAG 1171
1621 AGGAA 1625
1172 GGCAA 1176

RESULT 12
US-10-188-186-83
; Sequence 83, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 83
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1189)
US-10-188-186-83

Query Match 48.3%; Score 1130.6; DB 12; Length 1189;
Best Local Similarity 99.2%; Pred. No. 9.2e-301;
Matches 1136; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 481 AGAAGCACACTTTGTGGCCAAATAATGGTGTTCGCAATTTCTGCTGCTGATGGCC 540
Db 32 AGAAGCACACTTTGTGGCCAAATAATGGTGTTCGCAATTTCTGCTGCTGATGGCC 91
Qy 541 TGCCTGCTCCAGGAGAGAGCTTTGAAATGCTCATTTGTGGAGAGCTTCATCATGGGCATA 600
Db 92 TGCCTGCTCCAGGAGAGAGCTTTGAAATGCTCATTTGTGGAGAGCTTCATCATGGGCATA 151
Qy 601 GATGGAGCGCTCGCCCTCAGTGTCTCCCATGTACCTCAGTGTGAGATCTCACCAAGGAG 660
Db 152 GATGGAGCGCTCGCCCTCAGTGTCTCCCATGTACCTCAGTGTGAGATCTCACCAAGGAG 211
Qy 661 ATCCGTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACATGGG 720
Db 212 ATCCGTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACATGGG 271
Qy 721 CAGCTTCTGGGCTGCTGGGAGAGAGTACCTGGCCATACCTGTTTGGGA 780
Db 272 CAGCTTCTGGGCTGCTGGGAGAGAGTACCTGGCCATACCTGTTTGGGA 331
Qy 781 GTGATTTGTGCTCCGCTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCCGGAGAGCCCA 840
Db 332 GTGATTTGTGCTCCGCTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCCGGAGAGCCCA 391
Qy 841 CGCTACCTGCTCTTGGAGAGAGCAACAGAGGAGAGCTGTGAAAGCTTTCCAAACGTTTC 900
Db 392 CGCTACCTGCTCTTGGAGAGAGCAACAGAGGAGAGCTGTGAAAGCTTTCCAAACGTTTC 451
Qy 901 TTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAGGTCTCTGGCTGAGAGCCACGTGCAG 960
Db 452 TTGGGTAAAGCAGACATTTCCCAAGAGGTAGAGAGGTCTCTGGCTGAGAGCCACGTGCAG 511
Qy 961 AGGAGCATCCGCTGGTGTCCGCTGCTGAGAGCTGCTGAGAGCTCCCTACGTCGCTGCGCAG 1020
Db 512 AGGAGCATCCGCTGGTGTCCGCTGCTGAGAGCTGCTGAGAGCTCCCTACGTCGCTGCGCAG 571
Qy 1021 GTGTCACCGTGAATGTCAATGGCCCTGCTACAGCTCTGTGGCCTCAATGCAATTTGG 1080
Db 572 GTGTCACCGTGAATGTCAATGGCCCTGCTACAGCTCTGTGGCCTCAATGCAATTTGG 631
Qy 1081 TTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 1140
Db 632 TTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 691
Qy 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTCTGTTTGGTCAATTGAG 1200
Db 692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTCTGTTTGGTCAATTGAG 751
Qy 1201 CACCTGGAGCGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260
Db 752 CACCTGGAGCGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 811

Db 337 CTCTGGTCTGTGACTGTCCATATTCGCCATCGGTGGACTTGTGGGACGTTAAATGTG 396
QY 457 AAGATGATTGGAAGGTTCTTGGAGGAGACACACTTTGCTGCCAATATGTTTGA 516
Db 397 AAGATGATTGGAAGGTTCTTGGAGGAGACACACTTTGCTGCCAATATGTTTGA 456
QY 517 ATTTCTGCTGATGCTGATGGCTGCTCGCTCCAGGAGGAGCTTTGAAATGCTCAAT 576
Db 457 ATTTCTGCTGATGCTGATGGCTGCTCGCTCCAGGAGGAGCTTTGAAATGCTCAAT 516
QY 577 GTGGAGGCTTTCAATGAGGATAGATGGAGGCTGCTCCCTCAGTGTGCTCCCATGTAC 636
Db 517 GTGGAGGCTTTCAATGAGGATAGATGGAGGCTGCTCCCTCAGTGTGCTCCCATGTAC 576
QY 637 CTCAGTGAGATCTCACCAAGGATCGGTGGCTCTCTGGGAGGAGTGTGCTCAATCTTT 696
Db 577 CTCAGTGAGATCTCACCAAGGATCGGTGGCTCTCTGGGAGGAGTGTGCTCAATCTTT 636
QY 697 ATCTGCAATGGCGTGTTCACCTGGGAGCTTCTGGGCTGCTGGGAGGAG 756
Db 637 ATCTGCAATGGCGTGTTCACCTGGGAGCTTCTGGGCTGCTGGGAGGAG 696
QY 757 AGTACCTGGCCATACCTGTTTGGAGTGAATGTGCTCCCTGCTGCTGCTGAGC 816
Db 697 AGTACCTGGCCATACCTGTTTGGAGTGAATGTGCTCCCTGCTGCTGAGC 756
QY 817 CTTCCTTTCTCCCGGACAGCCACGCTACCTGCTCTTGAGAGACACACAGGCAAGA 876
Db 757 CTTCCTTTCTCCCGGACAGCCACGCTACCTGCTCTTGAGAGACACACAGGCAAGA 816
QY 877 GCTGTGAAGCCTTCCAAACGCTTCTGGGTAAAGCAGAGCTTTCCCAAGAGTAGAGG 936
Db 817 GCTGTGAAGCCTTCCAAACGCTTCTGGGTAAAGCAGAGCTTTCCCAAGAGTAGAGG 876
QY 937 GTCCTGCTGAGAGGACAGCTGACAGGAGCATCGCTGCTGCTGCTGCTGCTGCTG 996
Db 877 GTCCTGCTGAGAGGACAGCTGACAGGAGCATCGCTGCTGCTGCTGCTGCTGCTG 936
QY 997 AGAGTCCCTACGTCGCTGCGAGTGTGTCACCTGATTTGTCACATGGGCTGCTACAG 1056
Db 937 AGAGTCCCTACGTCGCTGCGAGTGTGTCACCTGATTTGTCACATGGGCTGCTACAG 996
QY 1057 CTCTGTGGCTTCAATGCAATTTGTTCTATACCAAGAGCATTTTGGAAAGCTGGATC 1116
Db 937 CTCTGTGGCTTCAATGCAATTTGTTCTATACCAAGAGCATTTTGGAAAGCTGGATC 1056
QY 1117 CCTCGGCAAGATCCCATACGTCACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCC 1176
Db 1057 CCTCGGCAAGATCCCATACGTCACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCC 1116
QY 1177 GTCCTCTGCTGTTGGTCAATGAGCAGCTGGAGAGACCCCTCCTCATTTGGTGGCTTT 1236
Db 1117 GTCCTCTGCTGTTGGTCAATGAGCAGCTGGAGAGACCCCTCCTCATTTGGTGGCTTT 1176
QY 1237 GGGCTCATGGGCTCTTTCTTTGGAGCCCTCACCATCAGCTGACCCCTGCGAGGACCGCC 1296
Db 1177 GGGCTCATGGGCTCTTTCTTTGGAGCCCTCACCATCAGCTGACCCCTGCGAGGACCGCC 1236
QY 1297 CCCTGGGTCCCTACCTGAGTATCGTGGGCAATTCGTGGCCATCATCGCTCTTTCTGAGT 1356
Db 1237 CCCTGGGTCCCTACCTGAGTATCGTGGGCAATTCGTGGCCATCATCGCTCTTTCTGAGT 1296
QY 1357 GGGCCAGGTGGATCCC 1373
Db 1297 GGGCCAGGTGTTTCCC 1313

RESULT 11
US-10-188-186-85
; Sequence 85, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:

; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO: 85
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1189)
US-10-188-186-85

Query Match 48.4%; Score 1133.8; DB 12; Length 1189;
Best Local Similarity 99.4%; Pred. No. 1.2e-301;
Matches 1138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 481 AGGAAGACACATTTGCTGGCCAAATATGGGTTTGCAATTTCTGTGCAATGCTGATGGCC 540
Db 32 AGGAAGACACATTTGCTGGCCAAATATGGGTTTGCAATTTCTGTGCAATGCTGATGGCC 91
QY 541 TGCTCGCTCCAGGAGGAGCCCTTGAATGCTCATTTGGGACGCTTCAATGAGGCATA 600
Db 92 TGCTCGCTCCAGGAGGAGCCCTTGAATGCTCATTTGGGACGCTTCAATGAGGCATA 151
QY 601 GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 660
Db 152 GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 211
QY 661 ATCCGTGGCTCTCTGGGAGGAGTGTGCTGCTATTTATCTGCAATTTGGGCTGTTCACTGGG 720
Db 212 ATCCGTGGCTCTCTGGGAGGAGTGTGCTGCTATTTATCTGCAATTTGGGCTGTTCACTGGG 271
QY 721 CAGCTTTCTGGGCTGCCCCAGCTGCTGGGAAAGGAGAGTACCTGCGCATACCTGTTTGA 780
Db 272 CAGCTTTCTGGGCTGCCCCAGCTGCTGGGAAAGGAGAGTACCTGCGCATACCTGTTTGA 331
QY 781 GTGATTTGGTCCCTGCGCTTGTCCAGTGTGAGGCTTCCCTTTCTCCGAGCAGCCCA 840
Db 332 GTGATTTGGTCCCTGCGCTTGTCCAGTGTGAGGCTTCCCTTTCTCCGAGCAGCCCA 391
QY 841 CGCTACCTGCTCTTGGAGNAGCACACAGGAGGAGTGTGAAAGCCCTTCCAAAGCTTC 900
Db 392 CGCTACCTGCTCTTGGAGNAGCACACAGGAGGAGTGTGAAAGCCCTTCCAAAGCTTC 451
QY 901 TTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCGAG 960
Db 452 TTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCGAG 511

1081	Db	CACCTTGAGTACAGGGGGCATCGAGACTTTGGCTCGGCTCTCTC-----	1125
1200	Qy	GCACCTGGACGAGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGG	1259
1126	Db	-----	1125
1260	Qy	GACCTCACCATCA CGCTGACCCCTGCAGGACACGCGCCCTGGGTCCCTACCTGAGTAT	1319
1126	Db	-----	1125
1320	Qy	CGTGGGCATTCTGGCCATCATGCGCTCTTTCTGCACTGGCGCAGGTGGCATCCCGTTCA	1379
1126	Db	-----TGGCATCCCGTTCA	1140
1380	Qy	CTTGACTGTGAGTTCTTCCAGCAATCTCAGCGCGCGGTGCCTTCATCATTTGCAAGCAC	1439
1141	Db	CTTGACTGTGAGTTCTTCCAGCAATCTCAGCGCGCGGTGCCTTCATCATTTGCAAGCAC	1200
1440	Qy	CGTCAACTGGCTCTCCAACTTTTGGCTGTTGGGCTCCTCTTCCCAATCATTCAGAAAAGTCT	1499
1201	Db	CGTCAACTGGCTCTCCAACTTTTGGCTGTTGGGCTCCTCTTCCCAATCATTCAGAAAAGTCT	1260
1500	Qy	GGACACCTACTGTTTCTTAGTCTTTGGCTACAAATTTGTATCAGAGTGCTATCTACTGTA	1559
1261	Db	GGACACCTACTGTTTCTTAGTCTTTGGCTACAAATTTGTATCAGAGTGCTATCTACTGTA	1320
1560	Qy	TTTTGTGCTGCTTGAGACCAAAAAAGAACCTTATGCAGAAATCAGCAGGCCATTTTCCAA	1619
1321	Db	TTTTGTGCTGCTTGAGACCAAAAAAGAACCTTATGCAGAAATCAGCAGGCCATTTTCCAA	1380
1620	Qy	AAGGAACAAAGCATACCCACCAAGAGAGAAAAATCGACTCAGCTGTGCTCATGTCT	1674
1381	Db	AAGGAACAAAGCATACCCACCAAGAGAGAAAAATCGACTCAGCTGTGCTCATGTGCT	1435

```

RESULT 9
US-10-188-186-89
; Sequence 89, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALA
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 89
; LENGTH: 1267
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

Db 241 TCACAGGTGGGCAAGCTTAGCCGCCAGAAAAGTCAAGTCACCTGCTCAGACTCCTCAGCT 300
QY 337 -----TACAATAGATCATGGGAAAGAGCATGGACGTCCAATAGAGCCAGACACTCTGA 391
Db 301 GAGGGGACTGGCCCTGGAGGTAAAGCTGATATCATCTTGGCTCAAAGCCCAAGCTCTAT 360
QY 392 CTCTCTCTGG-----TCTGTGACTGTGTCCATTAATTCGCC 426
Db 361 CTCGTGGCTGGTGGCACTAGAGGAGACAAACGAGATTGGCAGAGACTGGTCTCTGCTGGT 420
QY 427 ATCGGTGGACATTGTGGGACCTTAATTGTG-----AAGATGAT 464
Db 421 CCTGTGGCTCCCTCGCGGGCGCTTCGGCTCCTCTCTACCGCTACAACTGTGC 480
QY 465 TGGAAAGGTCTTGGGAGGAAGCACACTTTGCTGGCCCAATAATAGGTTTGCATTTCTGC 524
Db 481 GGTGTGAATGCCGCCCAACCCCGCACACTTTGCTGGCCCAATAATAGGTTTGCATTTCTGC 540
QY 525 TSCATTGTGATGGCTCTCGCTCCAGGAGGAGCTTTGAATGCTCATTTGTGGAGG 584
Db 541 TGCATTGTGATGGCTCTCGCTCCAGGAGGAGCTTTGAATGCTCATGTGGGAGG 600
QY 585 CTTTCATCATGGGCATAGATGAGGCGTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGA 644
Db 601 CTTTCATCATGGGCATAGATGAGGCGTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGA 660
QY 645 GATCTCACCAAGGAGATCCGTGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAT 704
Db 661 GATCTCACCAAGGAGATCCGTGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAT 720
QY 705 TGGCGTGTCTACTGGGAGCTTCTGGGCTGCCGAGCTGCTGGGAAAGAGAGTACTTG 764
Db 721 TGGCGTGTCTACTGGGAGCTTCTGGGCTGCCGAGCTGCTGGGAAAGAGAGTACTTG 780
QY 765 GCCATACCTGTTTGGAGTGATTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
Db 781 GCCATACCTGTTTGGAGTGATTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 825 TCTCCCGAGACCCACCTACTGCTCTCTGGAGAAGCACAACAGGAGCAAGACTGTGAA 884
Db 841 TCTCTGGACACCCACCTACTGCTCTCTGGAGAAGCACAACAGGAGCAAGACTGTGAA 900
QY 885 AGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTTCCAGAGGTAGAGAGTCTCTGCG 944
Db 901 AGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTTCCAGAGGTAGAGAGTCTCTGCG 960
QY 945 TGAGAGCCACTGCAGAGGACATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
Db 961 TGAGAGCCCGCTGCAGAGGACATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1005 CTAGCTCCGCTGGCAGGTGGTCAACCGTGAATGTCAACCATGCGCTGCTGCTGCTGCTGCTGCTGCT 1064
Db 1021 CTAGCTCCGCTGGCAGGTGGTCAACCGTGAATGTCAACCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1065 CCTCAATGCAATTTGGTTCTATACCAAGCAGATCTTTGGAAAAGCTGGGATCCCTCGGC 1124
Db 1081 CCTCAATGCAATTTGGTTCTATACCAAGCAGATCTTTGGAAAAGCTGGGATCCCTCGGC 1140
QY 1125 AAAGATCCCATACGTCACTTGGTACAGGGGGCATCGAGACTTTGGCTCCGCTCTCTCTC 1184
Db 1141 AAAGATCCCATACGTCACTTGGTACAGGGGGCATCGAGACTTTGGCTCCGCTCTCTCTC 1200
QY 1185 TGGTTTGGTCAATGAGCACTTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCAT 1244
Db 1201 TGGTTTGGTCAATGAGCACTTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCAT 1260
QY 1245 GGGCTCTCTTTGGGACCTTCAACATCAAGCTGACCTGACCTGAGGACCAAGCCCTCGGCT 1304
Db 1261 GGGCTCTCTTTGGGACCTTCAACATCAAGCTGACCTGACCTGAGGACCAAGCCCTCGGCT 1320
QY 1305 CCCCTTACCTGATGCTGGGCACTTCTGGCCCATCATGCTCTCTTTGCAAGTGGGCGAG 1364
Db 1321 CCCCTTACCTGATGCTGGGCACTTCTGGCCCATCATGCTCTCTTTGCAAGTGGGCGAG 1380

QY 1365 TGGCATCCCGTTCACTTTGACTGGTGAAGTTCTTCCAGCAATCTCAGCGCGGCTGCTT 1424
Db 1381 TGGCATCCCGTTCACTTTGACTGGTGAAGTTCTTCCAGCAATCTCAGCGCGGCTGCTT 1440
QY 1425 CATCATTCAGGACCGCTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCAT 1484
Db 1441 CATCATTCAGGACCGCTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCAT 1500
QY 1485 CATTCAGAAAAGTCTGGACACCTACTGTTTCTTAGTCTTTTGTCTACAAATTTGTATCACAG 1544
Db 1501 CATTCAGAAAAGTCTGGACACCTACTGTTTCTTAGTCTTTTGTCTACAAATTTGTATCACAG 1560
QY 1545 TGCTATCTACTGTTATTTTGTGCTGCTGAGACCAAAACAGAACCTATTCAGAAATCAG 1604
Db 1561 TGCTATCTACTGTTATTTTGTGCTGCTGAGACCAAAACAGAACCTATTCAGAAATCAG 1620
QY 1605 CCAGGCAATTTCCAAAAGGAAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGT 1664
Db 1621 CCAGGCAATTTCCAAAAGGAAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGT 1680
QY 1665 CACTGATGCT 1674
Db 1681 CACTGATGCT 1690

RESULT 8

US-10-188-186-79
; Sequence 79, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Angerson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 79
; LENGTH: 1502
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(1454)
US-10-188-186-79

Query Match 52.5%; Score 1229; DB 12; Length 1502;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 1425; Conservative 0; Mismatches 10; Indels 180; Gaps 1;
QY 60 GTCACTGAGACCCATGCGAGGAAACAAAATAGGAATTCAGAGAACTGGGCTAGTTCC 119

Qy 482 GGAAGCAGCTTTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGATTTGCTGATGGCT 541
Db 358 GGAAGCAGCTTTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGATTTGCTGATGGCT 417
Qy 542 GCTCGCTCCAGGAGAGAGCTTTGAATGCTATTGTTGGAGCGCTTCATCATGGCCATAG 601
Db 418 GCTCGCTCCAGGAGAGAGAGCTTTGAATGCTATTGTTGGAGCGCTTCATCATGGCCATAG 477
Qy 602 ATGAGAGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGTGAGATCTCACCAGAGGAGA 661
Db 478 ATGAGAGCGTGGCCCTCAGTGTGCTCCCATGTACCTTAGTGAGATCTCACCAGAGGAGA 537
Qy 662 TCGTGTGCTCTTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCATCTGGC 721
Db 538 TCGTGTGCTCTTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCATCTGGC 597
Qy 722 AGCTTCTGGGCTCGCCGAGCTGCTGGAAAGAGAGATACCTGGCCATACCTGTTTGGAG 781
Db 598 AGCTTCTGGGCTCGCCGAGCTGCTGGAAAGAGAGATACCTGGCCATACCTGTTTGGAG 657
Qy 782 TGATTTGGTCCCTGCGGTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCGGACAGCCAC 841
Db 658 TGATTTGGTCCCTGCGGTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCGGACAGCCAC 717
Qy 842 GCTACCTGCTCTGGAGAGACAACAGAGCAAGAGCTGTGAAGCCTTCCAAAGCTTCT 901
Db 718 GCTACCTGCTCTGGAGAGACAACAGAGCAAGAGCTGTGAAGCCTTCCAAAGCTTCT 777
Qy 902 TGGGTAAGCAGACGTTTCCAGAGAGTAGAGAGCTGCTGAGAGCCTCCCTACGTCGCTGGCAGG 961
Db 778 TGGGTAAGCAGACGTTTCCAGAGAGTAGAGAGCTGCTGAGAGCCTCCCTACGTCGCTGGCAGG 837
Qy 962 GGAGCATCCGCTGGTTCGCTGCTGAGCTGCTGAGAGCTCCCTACGTCGCTGGCAGG 1021
Db 838 GGAGCATCCGCTGGTTCGCTGCTGAGCTGCTGAGAGCTCCCTACGTCGCTGGCAGG 897
Qy 1022 TGGTACCGTGATTTGACCATGGCCCTGCTACAGCTCTGGGCTCCTCAATGCAATTTGGT 1081
Db 898 TGGTACCGTGATTTGACCATGGCCCTGCTACAGCTCTGGGCTCCTCAATGCAATTTGGT 957
Qy 1082 TCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCATACGTCA 1141
Db 958 TCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCATACGTCA 1017
Qy 1142 CCTTGAGTACAGGGGCGATCGAGACTTTGGCTGCGCTCTTCTCTGGTTTGGTCAATGAGC 1201
Db 1018 CCTTGAGTACAGGGGCGATCGAGACTTTGGCTGCGCTCTTCTCTGGTTTGGTCAATGAGC 1077
Qy 1202 ACCTGGAGCGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGA 1261
Db 1078 ACCTGGAGCGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGA 1137
Qy 1262 CCTTACCATACGCTGACCCCTGAGAGACACGCCCCCTGGGTCCCTTACCTGAGTATCG 1321
Db 1138 CCTTACCATACGCTGACCCCTGAGAGACACGCCCCCTGGGTCCCTTACCTGAGTATCG 1197
Qy 1322 TGGGCAATCTGGGCATCATGCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCACT 1381
Db 1198 TGGGCAATCTGGGCATCATGCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCACT 1257
Qy 1382 TGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGAGGAGCAGC 1441
Db 1258 TGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGAGGAGCAGC 1317
Qy 1442 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCATTCATTCAGAAAGTCTGG 1501
Db 1318 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCATTCATTCAGAAAGTCTGG 1377
Qy 1502 ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGATACAGGCTGCTATCTACCTGTATT 1561
Db 1378 ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGATACAGGCTGCTATCTACCTGTATT 1437

Qy 1562 TTGTGCTGCTGAGACCAAAAAACAGAACCTATATGCAAAATCAGCCAGGCAATTTTCCAAAA 1621
Db 1438 TTGTGCTGCTGAGACCAAAAAACAGAACCTATATGCAAAATCAGCCAGGCAATTTTCCAAAA 1497
Qy 1622 GGAACAAAGCATATCCCAACAGAGAGAAATCGACTGCTGCTACTGATGCT 1674
Db 1498 GGAACAAAGCATATCCCAACAGAGAGAAATCGACTGCTGCTACTGATGCT 1550

RESULT 7
US-10-188-186-77
; Sequence 77, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 77
; LENGTH: 1757
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(1709)
; US-10-188-186-77

Query Match 52.6%; Score 1232.8; DB 12; Length 1757;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1448; Conservative 0; Mismatches 167; Indels 75; Gaps 4;
Qy 60 GTCACTGAGACCCATGGCAAGGAAACAAATAGGAATTTCCAGGAACCTGGGCTGATTTCC 119
Db 1 GTCACTGAGACCCATGGCAAGGAAACAAATAGGAATTTCCAGGAACCTGGGCTGATTTCC 60
Qy 120 CCTCAGATGACACACGACCCGCGGCTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGG 179
Db 61 CCTCAGATGACACACGACCCGCGGCTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 180 TGACCACCTGAGGAGTGGGGTGGCCAGGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 234
Db 121 TGACCACCTGAGGAGTGGGGTGGCCAGGTGGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 235 TGGCTTCTGCTGGGCTCCCTTCCGCGGCGGCTTGGGCTCCTCTTCTTCTTCTTCTTCTTCT 294
Db 181 CTCCTCTGCGAGGCTCCTCAACAAACATATGTGCCAGTGTCTGTCTTATTAGATCCCAT 240
Qy 295 CTGTGGTGGTCAATGTCCTCCACCCCGTACATCAAGGCTTT----- 336

418	Db	GCTCGCTCCAGCAGGACGCTTTGAAATGCTCTCATCTGTGGGACGCTTONTCAATGAGGCATAG	477
602	Qy	ATGAGGCGTGCCTCCTCAGTGTGCTCCCATATGATACCTCAGTGAGATCTACCCCAAGGAGA	661
478	Db	ATGAGGCGTGCCTCCTCAGTGTGCTCCCATATGATACCTCAGTGAGATCTACCCCAAGGAGA	537
662	Qy	TCCGTGGCTCTCTGGGCGAGTGACTGCATCTTTATCTGCAATTTGGGTGTCTACTGGC	721
538	Db	TCCGTGGCTCTCTGGGCGAGTGACTGCATCTTTATCTGCAATTTGGGTGTCTACTGGC	597
722	Qy	AGCTTCTGGGCGCTGCCCGAGTCTGTGGAAAGGAGAGTAACCTGGCCATACCTGTTTGGAG	781
598	Db	AGCTTCTGGGCGCTGCCCGAGTCTGTGGAAAGGAGAGTAACCTGGCCATACCTGTTTGGAG	657
782	Qy	TGATTGTGGTCCCTTCGCGCTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGACAGCCAC	841
658	Db	TGATTGTGGTCCCTTCGCGCTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGACAGCCAC	717
842	Qy	GCTACCTGCTCTTGGAGAAGCACAAACAGGCAAGAGCTGTGAAAGCCTTCCAAACGTTCT	901
718	Db	GCTACCTGCTCTTGGAGAAGCACAAACAGGCAAGAGCTGTGAAAGCCTTCCAAACGTTCT	777
902	Qy	TGGTAAAGCAGACGTTTCCCAAGAGGTAGAGGAGTCTTGGCTGAGAGCCACGTCGAGA	961
778	Db	TGGTAAAGCAGACGTTTCCCAAGAGGTAGAGGAGTCTTGGCTGAGAGCCGCGTGCAGA	837
962	Qy	GGAGCATCCGCGCTGGTTCGCTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCGAGG	1021
838	Db	GGAGCATCCGCGCTGGTTCGCTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCGAGG	897
1022	Qy	TGCTCACCGTGATTGTACCATGCGCTGTACAGCTCTGTGGCGCTCAATGCAATTTGGT	1081
898	Db	TGCTCACCGTGATTGTACCATGCGCTGTACAGCTCTGTGGCGCTCAATGCAATTTGGT	957
1082	Qy	TCTATACCAACAGCATCTTTGGAAAAGCTGGATCCCTCCGGCAAAAGATCCCATACGTCA	1141
958	Db	TCTATACCAACAGCATCTTTGGAAAAGCTGGATCCCTCCGGCAAAAGATCCCATACGTCA	1017
1142	Qy	CTTTGAGTACAGGGGGCATACAGACTTTGGCTGCGCTTCTCTCTGTTTGGTCATTGAGC	1201
1018	Db	CTTTGAGTACAGGGGGCATACAGACTTTGGCTGCGCTTCTCTCTGTTTGGTCATTGAGC	1077
1202	Qy	ACCTGGACCGAGACCCCTCCTCATTTGTTGGCTTGGGCTCATGGGCTCTTTCTTTGGGA	1261
1078	Db	ACCTGGACCGAGACCCCTCCTCATTTGTTGGCTTGGGCTCATGGGCTCTTTCTTTGGGA	1137
1262	Qy	CCCTCACATCAGCTGACCTGCAGAACACGCCCCCTGGGTCCCTACCTCAGTATCG	1321
1138	Db	CCCTCACATCAGCTGACCTGCAGAACACGCCCCCTGGGTCCCTACCTCAGTATCG	1197
1322	Qy	TGGGCATCTGGGCCATCATCGCTCTTTCTGCACTGGGCCAGGTGGCATCCCCGTTCATCT	1381
1198	Db	TGGGCATCTGGGCCATCATCGCTCTTTCTGCACTGGGCCAGGTGGCATCCCCGTTCATCT	1257
1382	Qy	TGACTGTGTGAGTCTTCCAGCAATCTCAGCGCGCGCTGCTTCATCATTTGCAAGGCTGG	1441
1258	Db	TGACTGTGTGAGTCTTCCAGCAATCTCAGCGCGCGCTGCTTCATCATTTGCAAGGCTGG	1317
1442	Qy	TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCATTCATTTGAGAAAGTCTGG	1501
1318	Db	TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCATTCATTTGAGAAAGTCTGG	1377
1502	Qy	ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGTATACAGGTGCTATCTACCTGTATT	1561
1378	Db	ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGTATACAGGTGCTATCTACCTGTATT	1437
1562	Qy	TTGTGCTGCTGAGACCAAAAACAGAACCTTATCGAGAAATCAGCCAGGCATTTTCCAAA	1621
1438	Db	TTGTGCTGCTGAGACCAAAAACAGAACCTTATCGAGAAATCAGCCAGGCATTTTCCAAA	1497
1622	Qy	GGAAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTCATGTGCT	1674
1498	Db	GGAAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTCATGTGCT	1550

RESULT 6

```

US-10-117-722-574
; Sequence 574, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744
; TITLE OF INVENTION: Polyptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 574
; LENGTH: 2842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (345)..(1571)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2842)
; OTHER INFORMATION: n = a,t,c or g
US-10-117-722-574

```

Query Match	56.2%	Score	1317.8	DB	15	Length	2842
Best Local Similarity	89.9%	Pred. No.	0				
Matches 1504; Conservative	0	Mismatches	7	Indels	162	Gaps	2

QY	2	CGACCCACGGTCGCGGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGT	61
DB	40	CGACCCACGGTGC-CCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGT	98
QY	62	CACGTAGACCCATGGCAAGGAAACAAATAATAGGAATTCCAAGGAACTCGGGCTAGCTCCCC	121
DB	99	CACGTAGACCCATGGCAAGGAAACAAATAATAGGAATTCCAAGGAACTCGGGCTAGCTCCCC	158
QY	122	TCACAGATCACACGCCACGCCGGGCTCCAGGGGCCAGGAGGACACTGCTGGAGTGTG	181
DB	159	TCACAGATCACACCCAGCCAGCCAGGCTCCAGGGCCAGGAGGACACTGCTGGAGTGTG	218
QY	182	ACCACCTGAGGAGTGGGTGCCAGGTGGAGGAGAGAAAGACCTGCTCTGCTGCTCC	241
DB	219	ACCACCTGAGGAGTGGGTGCCAGGTGGAGGAGAGAAAGGACCTGCTCTGCTGCTCC	278
QY	242	TGTTGGCTTCCCTCGGGGGGCTTCGGTTCCTTCCTCTACGGGTACAACCTGTGCG	301
DB	279	TGTTGGCTTCCCTCGGGGGGCTTCGGTTCCTTCCTCTACGGGTACAACCTGTGCG	338
QY	302	TGTTGAATGCCCCACCCCGCTACATCAAGGCTTTTACAATGAGTCATGGAAAGAGGC	361
DB	339	TGTTGAATGCCCCACCC-----	357
QY	362	ATGGAGCTCCAATAGACCCAGACACTCTGACTGTGCTGCTGCTGTGTCATAT	421
DB	358	-----	357
QY	422	TCGCCATCGGTGGACTTGTGGGACGTAAATTGTGAAGATGATTTGGAAGGTTCTTGGGA	481
DB	358	-----	357

Db 622 AGTGATTGGTCCCTGCGGTTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGCCC 681
QY 840 AGCTACCTGCTCTTGGAGAAGCACAACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTT 899
Db 682 AGCTACCTGCTCTTGGAGAAGCACAACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTT 741
QY 900 CTTGGTAAAGCAGAGCTTTCCAGAGGTAGAGAGTCTCGCTGAGAGCCAGCTGCA 959
Db 742 CTTGGTAAAGCAGAGCTTTCCAGAGGTAGAGAGTCTCGCTGAGAGCCAGCTGCA 801
QY 960 GAGGAGCATCCGCTGGTGTCCGCTGTGAGAGTCTGAGAGTCCCTACGCTCCGCTGGCA 1019
Db 802 GAGGAGCATCCGCTGGTGTCCGCTGTGAGAGTCTGAGAGTCCCTACGCTCCGCTGGCA 861
QY 1020 GTTGTTACCGGTATGTACCATGGCTGTACAGCTCTGTGAGCTCTCAATGCAATTTG 1079
Db 862 GTTGTTACCGGTATGTACCATGGCTGTACAGCTCTGTGAGCTCTCAATGCAATTTG 921
QY 1080 GTTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCGGCAAAAGATCCCATAGT 1139
Db 922 GTTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCGGCAAAAGATCCCATAGT 981
QY 1140 CACCTTGTGTACAGGGGATCGAGACTTTGGCTGCGCTCTCTCTGTTGGTCAATGA 1199
Db 982 CACCTTGTGTACAGGGGATCGAGACTTTGGCTGCGCTCTCTCTGTTGGTCAATGA 1041
QY 1200 GCACCTGGAGCAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGG 1259
Db 1042 GCACCTGGAGCAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGG 1101
QY 1260 GACCTTACCATACGCTGACCTGAGGACCAACGCGGCTGGGTCCCTACCTGAGTAT 1319
Db 1102 GACCTTACCATACGCTGACCTGAGGACCAACGCGGCTGGGTCCCTACCTGAGTAT 1161
QY 1320 CGTGGGCTTGGGCTATCGCTCTCTTCTGAGTGGGCGAGTGGGCTCCGCTTCAT 1379
Db 1162 CGTGGGCTTGGGCTATCGCTCTCTTCTGAGTGGGCGAGTGGGCTCCGCTTCAT 1221
QY 1380 CTTGACTGTGTAGTTCTTCCAGCAATCTCAGGCGCGGTGCTCTCATCTTGCAGGCAC 1439
Db 1222 CTTGACTGTGTAGTTCTTCCAGCAATCTCAGGCGCGGTGCTCTCATCTTGCAGGCAC 1281
QY 1440 CGTCAACTGGCTCTCCAACTTTGGCTGTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCT 1499
Db 1282 CGTCAACTGGCTCTCCAACTTTGGCTGTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCT 1341
QY 1500 GAGACCTACTGTTTCCTAGTCTTTGCTACAAATTTGTATCAGGTGTATCTACCTGTA 1559
Db 1342 GAGACCTACTGTTTCCTAGTCTTTGCTACAAATTTGTATCAGGTGTATCTACCTGTA 1401
QY 1560 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCCAGGCAATTTCCAA 1619
Db 1402 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCCAGGCAATTTCCAA 1461
QY 1620 AAGGAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGCTGTCACTGATGCT 1674
Db 1462 AAGGAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGCTGTCACTGATGCT 1516

RESULT 5

US-10-037-270-574

; Sequence 574, Application US/10037270

; Publication No. US20030104529A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: Pt_FL_genes Version 1.0
; SEQ ID NO 574
; LENGTH: 2842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (345)..(1571)
; NAME/KEY: misc feature
; LOCATION: (1)..(2842)
; OTHER INFORMATION: n = a,t,c or g
US-10-037-270-574

Query Match 56.2%; Score 1317.8; DB 14; Length 2842;

Best Local Similarity 89.9%; Pred. No. 0;

Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;

QY 2 CGACCCACGGCTCCGGCTTTGGCAGAGTCTGGGGTCCCTGGAGTACAGCATCAGTGGGT 61
Db 40 CGACCCACGGCTGC-GCCTTTGGCAGAGTCTGGGGTCCCTGGAGTACAGCATCAGTGGGT 98
QY 62 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCGAAGGACTTGGGCTAGTCCCTCCC 121
Db 99 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCGAAGGACTTGGGCTAGTCCCTCCC 158
QY 122 TCACAGATGACACAGCAGCCGCGGCTCCAGGGCCAGGGAGGAGGACCTGCTGGAGTGTG 181
Db 159 TCACAGATGACACAGCAGCCGCGGCTCCAGGGCCAGGGAGGAGGACCTGCTGGAGTGTG 218
QY 182 ACCACCTGAGGAGTGGGGTCCAGAGTGGAGGAGAGAAAGGACCTGCTCGCTCC 241
Db 219 ACCACCTGAGGAGTGGGGTCCAGAGTGGAGGAGAGAAAGGACCTGCTCGCTCC 278
QY 242 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 301
Db 279 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 338
QY 302 TGGTGAATGCCCCCAGCCCGGTACATCAAGGCCCTTTTACATGAGTCAATGGGAAGAGGC 361
Db 339 TGGTGAATGCCCCCAGCCCGGTACATCAAGGCCCTTTTACATGAGTCAATGGGAAGAGGC 357
QY 362 ATGAGCTCCAAATAGACCCAGACACTCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
Db 358 ----- 357
QY 422 TCGCCATCGGTGGACTTGTGGGAGCGTTAATGTAAGATGATTTGAAAGGTTCTTGGGA 481
Db 358 ----- 357
QY 482 GGAAGCAGACTTTGCTGGCCCAATAATGGGTTTCCAAATTTCTGCTGATGCTGATGCTGCT 541
Db 358 GGAAGCAGACTTTGCTGGCCCAATAATGGGTTTCCAAATTTCTGCTGATGCTGATGCTGCT 417
QY 542 GCTCGCTCCAGGAGGAGGCTTTGAATGCTCAATTTCTGGAGCGCTTTCATCATCGGAGTAG 601

QY 984 GTGTGAGTGTGTGAGAGTCTCCCTAGTCCGTGGCAGGTGTGTCACCGTATGTGTACCAT 1043
Db 968 GCTGAGTGTGTGAGAGTCTCCCTAGTCCGTGGCAGGTGTGTCACCGTATGTGTACCAT 1027
QY 1044 GGCCTGTACAGTCTGTGGCTCAATGCAATTTGGTTCTATACCAAGCAATCTTTGG 1103
Db 1028 GGCCTGTACAGTCTGTGGCTCAATGCAATTTGGTTCTATACCAAGCAATCTTTGG 1087
QY 1104 AAAAGCTGGATCCCTCCGGCAAGATCCCATAGTCACTTGTGATACAGGGGGCATCGA 1163
Db 1088 AAAAGCTGGATCCCTCCGGCAAGATCCCATAGTCACTTGTGATACAGGGGGCATCGA 1147
QY 1164 GACTTGTGCTGCGTCTCTCTGTTTGGTCAATGAGCACTGGGAGAGCCCTCTCT 1223
Db 1148 GACTTGTGCTGCGTCTCTCTGTTTGGTCAATGAGCACTGGGAGAGCCCTCTCT 1207
QY 1224 CATTTGGTCTTGGGCTATGGGCTCTCTTTGGAGCCCTCACCATACGTCAGCCCT 1283
Db 1208 CATTTGGTCTTGGGCTATGGGCTCTCTTTGGAGCCCTCACCATACGTCAGCCCT 1267
QY 1284 GCAGGACCAACCCCTCTGGTCCCTTACCTGAGTATGTTGGGCAATCTGGCCATCATCGC 1343
Db 1268 GCAGGACCAACCCCTCTGGTCCCTTACCTGAGTATGTTGGGCAATCTGGCCATCATCGC 1327
QY 1344 CTCTTTGTGAGTGGGCGAGTGGCATCCCGTTTCATCTTGACTGGTGGTCTTCTCCAGCA 1403
Db 1328 CTCTTTGTGAGTGGGCGAGTGGCATCCCGTTTCATCTTGACTGGTGGTCTTCTCCAGCA 1387
QY 1404 ATCTCAGCGGGGCTGCTTCATCATTTGAGGACCGTCACTGGCTCTCCAACTTTGC 1463
Db 1388 ATCTCAGCGGGGCTGCTTCATCATTTGAGGACCGTCACTGGCTCTCCAACTTTGC 1447
QY 1464 TGTGGGCTCTCTCTTCCCATTCATTCAGAAAGTCTGGACACCTACTGTTTCTAGTCTT 1523
Db 1448 TGTGGGCTCTCTCTTCCCATTCATTCAGAAAGTCTGGACACCTACTGTTTCTAGTCTT 1507
QY 1524 TGCTACAAATTTGTATCAGTGTCTATCTACCTGTATTTTGTGCTGCTGAGACCAAAA 1583
Db 1508 TGCTACAAATTTGTATCAGTGTCTATCTACCTGTATTTTGTGCTGCTGAGACCAAAA 1567
QY 1584 CAGAACCTATGCAGAAATCAGCAGGCAATTTCCAAAGGAAACAAAGCATACCCACCA 1643
Db 1568 CAGAACCTATGCAGAAATCAGCAGGCAATTTCCAAAGGAAACAAAGCATACCCACCA 1627
QY 1644 AGAGAAATCGACTCAGCTGTCTGATGCT 1674
Db 1628 AGAGAAATCGACTCAGCTGTCTGATGCT 1658

RESULT 4

US-10-188-186-75
; Sequence 75, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502

; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 75
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(1535)
US-10-188-186-75

Query Match 59.7%; Score 1399; DB 12; Length 1710;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 5; Indels 99; Gaps 1;
QY 60 GTCACTGAGACCCATGCAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCC 119
Db 1 GTCACTGAGACCCATGCAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCC 60
QY 120 CCTCACAGATGACACGACGACGCGGGCTCCAGGCGCAGGAGGCGACTGCTGAGTG 179
Db 61 CCTCACAGATGACACGACGACGCGGGCTCCAGGCGCAGGAGGCGACTGCTGAGTG 120
QY 180 TGACCACTCTGAGGAGTGGGCTGCGAGGTGGAAGGAGAAAGCACTGGTCTGCTCGCT 239
Db 121 TGACCACTCTGAGGAGTGGGCTGCGAGGTGGAAGGAGAGAA----- 162
QY 240 CCTCGTGGCTCTCGTGGGCGGCGCTTCGGCTCCTCTCTACGGCTACAACTGTG 299
Db 163 ----- 162
QY 300 GGTGTGAATGCCCCACCCGCTACATCAAGGCTTTTACATGAGTCAATGGAAAGAG 359
Db 163 -----GTACATCAAGGCTTTTACATGAGTCAATGGAAAGAG 201
QY 360 GCATGAGACGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTGCT 419
Db 202 GCATGAGACGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTGCT 261
QY 420 ATTGCGCATCGTGGACTTGTGGGACGTTAATTGTGAAGATGATTTGGAAGAGTTCTTTGG 479
Db 262 ATTGCGCATCGTGGACTTGTGGGACATTAATTGTGAAGATGATTTGGAAGAGTTCTTTGG 321
QY 480 GAGGAGACACACTTTGCTGGCCCAATATCGGTTTGCATTTCTGCTGCAATTCCTGATGC 539
Db 322 GAGGAGACACACTTTGCTGGCCCAATATCGGTTTGCATTTCTGCTGCAATTCCTGATGC 381
QY 540 CTGCTGCTCCAGGACGAGGCTTTTGAATGCTCATTTGTGGGACGCTTCATCATGGCAT 599
Db 382 CTGCTGCTCCAGGACGAGGCTTTTGAATGCTCATTCGTEGGACGCTTCATCATGGCAT 441
QY 600 AGATGAGGCGCTGCGCTCAGTGTGCTCCCGCTACCTCTGCTGAGATCTCACCAAGAG 659
Db 442 AGATGAGGCGCTGCGCTCAGTGTGCTCCCGCTACCTCTGCTGAGATCTCACCAAGAG 501
QY 660 GATCCGTGGCTCTCTGGGCGAGGTGATCCCATCTTTATCTGCAATGGGCTTCACTGG 719
Db 502 GATCCGTGGCTCTCTGGGCGAGGTGATCCCATCTTTATCTGCAATGGGCTTCACTGG 561
QY 720 GCAGCTTCTGGGCTCTCCCGAGCTGCTGGAAAGGAGAGTACTCTGCCATACCTGTTGG 779
Db 562 GCAGCTTCTGGGCTCTCCCGAGCTGCTGGAAAGGAGAGTACTCTGCCATACCTGTTGG 621
QY 780 AGTGAATTTGCTCCCTGCGGCTTGTGCTGAGCTGCTGAGCTTCTCCCGGACAGCCC 839
|

Db	841	AGAAGCACAAAGGACAAGAGCTGTGAAGACGTTTCAAAACGTTCTTGGGTAAAGCAGACG	900
Qy	917	TTTCCCAAGAGGTAGAGGAGGTCCCTGGCTGAGAGGCCACGTCGAGAGGAGCATCCGCGCTGG	976
Db	901	TTTCCCAAGAGGTAGAGGAGGTCCCTGGCTGAGAGCGCGTGCAGAGGAGCATCCGCGCTGG	960
Qy	977	TGTCGCTGCTGGAGCTGCTGAGAGCTCCCTACCTCCGCTGGCAGGTGGTCAACCGTGATTG	1036
Db	961	TGTCGCTGCTGGAGCTGCTGAGAGCTCCCTACCTCCGCTGGCAGGTGGTCAACCGTGATTG	1020
Qy	1037	TCACCATGGCTGCTACAGAGCTCTGTGGCTCAATGCAATTTGGTCTTATACCAACAGCA	1096
Db	1021	TACCATGGCTGCTACAGAGCTCTGTGGCTCAATGCAATTTGGTCTTATACCAACAGCA	1080
Qy	1097	TCTTTGAAAAGCTGGGATCCCTCCGCAAGAATCCCATACGTCACCTTGGTGGTACAGGG	1156
Db	1081	TCTTTGAAAAGCTGGGATCCCTCTGGCAAGAATCCCATACGTCACCTTGGTGGTACAGGG	1140
Qy	1157	GCATCGAGACTTTGGCTGCCCTCTCTCTGTTTGGTCAATGAGACACCTGGGACGGAGAC	1216
Db	1141	GCATCGAGACTTTGGCTGCCCTCTCTCTGTTTGGTCAATGAGACACCTGGGACGGAGAC	1200
Qy	1217	CCCTGCTCATTTGGTGGCTTTGGGCTCATAGGGCTCTCTTTGGGACCTCAACATCAACG	1276
Db	1201	CCCTGCTCATTTGGTGGCTTTGGGCTCATAGGGCTCTCTTTGGGACCTCAACATCAACG	1260
Qy	1277	TGACCTCTCAGAACACAGCCCCCTGGTGCCTTACCTGAGTATCGTGGGACCTCTCGGCCA	1336
Db	1261	TGACCTCTCAGAACACAGCCCCCTGGTGCCTTACCTGAGTATCGTGGGACCTCTCGGCCA	1320
Qy	1337	TCATCGCCTCTTTCTGCAGTGGGCGAGGTGGCATCCCGTTTCATCTTGACTGGTCAAGTTCT	1396
Db	1321	TCATCGCCTCTTTCTGCAGTGGGCGAGGTGGCATCCCGTTTCATCTTGACTGGTCAAGTTCT	1380
Qy	1397	TCCAGCAATCTCAGCGCGCGGTGCCCTTCATCATTTGACGACGCAACGTCAGCTGGCTCTCCA	1456
Db	1381	TCCAGCAATCTCAGCGCGCGGTGCCCTTCATCATTTGACGACGCAACGTCAGCTGGCTCTCCA	1440
Qy	1457	ACTTTGCTGTTGGGCTCCTCTTCCCATTCATTCAGAAAAGCTGGAACACTACGTTGCTCC	1516
Db	1441	ACTTTGCTGTTGGGCTCCTCTTCCCATTCATTCAGAAAAGCTGGAACACTACGTTGCTCC	1500
Qy	1517	TAGTCTTTGCTACAAATTGATCACAGGTGCTATCTTACCTGTATTTTGGTGTGCTGCTGAGA	1576
Db	1501	TAGTCTTTGCTACAAATTGATCACAGGTGCTATCTTACCTGTATTTTGGTGTGCTGCTGAGA	1560
Qy	1577	CCAAAAACAGAACCTATGCAGAAATCAGCGAGGCATTTTCCAAAAGGAACAAAGCATACC	1636
Db	1561	CCAAAAACAGAACCTATGCAGAAATCAGCGAGGCATTTTCCAAAAGGAACAAAGCATACC	1620
Qy	1637	CACCAAGAGAAATTCGACTACGCTGTGCACTGATGCTC	1675
Db	1621	CACCAAGAGAAATTCGACTACGCTGTCACTGATGCTC	1659

```

RESULT 3
US-10-276-774-930
; Sequence 930, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom

```


LENGTH: 2343 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 73...1761
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-981-947A-1

Query Match 100.0%; Score 2343; DB 9; Length 2343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGACCCACGCTCCGGCCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG 60
Db
Qy 1 TCGACCCACGCTCCGGCCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG 60
Db
Qy 61 TCACTGAGACCCATGGCAGGAAACAAATAGAAATTCAGAGAACTGGGCCCTAGTCC 120
Db 61 TCACTGAGACCCATGGCAGGAAACAAATAGAAATTCAGAGAACTGGGCCCTAGTCC 120
Qy 121 CTCACAGATACACACGACCCAGCGCGGCTCCAGGCCCAAGGAGGCACTGCTGAGTGT 180
Db 121 CTCACAGATACACACGACCCAGCGCGGCTCCAGGCCCAAGGAGGCACTGCTGAGTGT 180
Qy 181 GACCACCTGAGGAGTGGGGTCCAGGTGGAAGAGAGAAAGACTGTGCTCGCTC 240
Db 181 GACCACCTGAGGAGTGGGGTCCAGGTGGAAGAGAGAAAGACTGTGCTCGCTC 240
Qy 241 CTCGTGCTCCCTCGCGGCGCTTCGGCTCCTCTCTACGGCTACAACTGTCG 300
Db 241 CTCGTGCTCCCTCGCGGCGCTTCGGCTCCTCTCTACGGCTACAACTGTCG 300
Qy 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTACATGAGTCAATGGAAAGAG 360
Db 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTACATGAGTCAATGGAAAGAG 360
Qy 361 CATGGAGCTCCATAGACCCAGACCTCTGACTCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 CATGGAGCTCCATAGACCCAGACCTCTGACTCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 TTGCCATCGGTGGACTGTGGGAGCTTAATGTGAAGATGATTTGGAAGGTTCTTGG 480
Db 421 TTGCCATCGGTGGACTGTGGGAGCTTAATGTGAAGATGATTTGGAAGGTTCTTGG 480
Qy 481 AGAAGACACTTTGCTGGCCAAATATGGGTTTGGAAATTTCTGCTGCTGCTGCTGCTG 540
Db 481 AGAAGACACTTTGCTGGCCAAATATGGGTTTGGAAATTTCTGCTGCTGCTGCTGCTG 540
Qy 541 TGCTCGCTCCAGCAGAGAGCTTTGAAATGCTCATTTGGGAGCGCTTCATCATGGGCAT 600
Db 541 TGCTCGCTCCAGCAGAGAGCTTTGAAATGCTCATTTGGGAGCGCTTCATCATGGGCAT 600
Qy 601 GATGGAGGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAGGAG 660
Db 601 GATGGAGGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAGGAG 660
Qy 661 ATCGTGGCTCTCTGGGCGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 ATCGTGGCTCTCTGGGCGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 CAGCTTCTGGGCTCGCCGAGCTGCTGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 780
Db 721 CAGCTTCTGGGCTCGCCGAGCTGCTGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 780
Qy 781 GTGATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 GTGATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 CGTACTGCTCTGTGAGAGACCAACAGGAGCAAGAGCTGTGAAAGCTTCCAAAGCTTC 900
Db

Db 841 CGCTACTGCTCTTTGGAGAAGCACACAGGAGCAAGAGCTGTGAAAGCTTCCAAAGCTTC 900
Qy 901 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTCGGCTGAGAGCCACCTGCGAG 960
Db 901 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTCGGCTGAGAGCCACCTGCGAG 960
Qy 961 AGGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 AGGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 GTGGTCACCGTGTATGTGTACCAATGGCCCTGCTACACAGCTCTGTGGCTCAATGCAATTTGG 1080
Db 1021 GTGGTCACCGTGTATGTGTACCAATGGCCCTGCTACACAGCTCTGTGGCTCAATGCAATTTGG 1080
Qy 1081 TTCTATACCAACAGCATCTTTTGGAAAGCTGGATCCCTCCGGCAAAAGATCCCATACCTC 1140
Db 1081 TTCTATACCAACAGCATCTTTTGGAAAGCTGGATCCCTCCGGCAAAAGATCCCATACCTC 1140
Qy 1141 ACCTTGAGTACAGGGGCAATCGAGACTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 ACCTTGAGTACAGGGGCAATCGAGACTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Qy 1201 CACCTGGAGCGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG 1260
Db 1201 CACCTGGAGCGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG 1260
Qy 1261 ACCCTCACCATCAACGCTGACCTGACGACACGAGCCCTCGGCTCCCTACTAGTATC 1320
Db 1261 ACCCTCACCATCAACGCTGACCTGACGACACGAGCCCTCGGCTCCCTACTAGTATC 1320
Qy 1321 GTGGGCAATTTCTGGCCATCATCGCTCTTTTGTGAGTGGGCGAGGTGGCATCCGCTTATC 1380
Db 1321 GTGGGCAATTTCTGGCCATCATCGCTCTTTTGTGAGTGGGCGAGGTGGCATCCGCTTATC 1380
Qy 1381 TTGACTGTGAGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGCAAGGAC 1440
Db 1381 TTGACTGTGAGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGCAAGGAC 1440
Qy 1441 GTCAACTGCTCTCCAACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1441 GTCAACTGCTCTCCAACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Qy 1501 GACACCTACTGTTTCTAGTCTTCTGCTCAATTTGTATCAAGGTGCTATCTAGCTGAT 1560
Db 1501 GACACCTACTGTTTCTAGTCTTCTGCTCAATTTGTATCAAGGTGCTATCTAGCTGAT 1560
Qy 1561 TTGTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGCAGGCAATTTTCCAA 1620
Db 1561 TTGTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGCAGGCAATTTTCCAA 1620
Qy 1621 AGAACAAGCATAACCCACAGAGAAATTCGACTCAGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1621 AGAACAAGCATAACCCACAGAGAAATTCGACTCAGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Qy 1681 TCTTCTCTTCTCACTTCCGAAATACGCTGGAATTCAGGTGCGGCAACCAACCGCC 1740
Db 1681 TCTTCTCTTCTCACTTCCGAAATACGCTGGAATTCAGGTGCGGCAACCAACCGCC 1740
Qy 1741 ACCAAAAGACACCCATTTGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1741 ACCAAAAGACACCCATTTGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Qy 1801 TCCCTTATCTTCTCTCTGAGAAACACAAAGTCTATGATGTGAGCAAGAGCTTGGATTT 1860
Db 1801 TCCCTTATCTTCTCTCTGAGAAACACAAAGTCTATGATGTGAGCAAGAGCTTGGATTT 1860
Qy 1861 TGGAGACATGGGTTTGAATTCAGTCTATCTCTTTTATTTAGCAAAATATTTAAACAGT 1920
Db 1861 TGGAGACATGGGTTTGAATTCAGTCTATCTCTTTTATTTAGCAAAATATTTAAACAGT 1920
Qy 1921 ACTGACATGCTCCATATGTTGTTTCCACCTGCTTATACAAATGGAGGAGAGAGAG 1980
Db 1921 ACTGACATGCTCCATATGTTGTTTCCACCTGCTTATACAAATGGAGGAGAGAGAG 1980

Sequence 39, Appl
Sequence 57, Appl
Sequence 696, App
Sequence 861, App
Sequence 307, App
Sequence 25, Appl
Sequence 47, Appl
Sequence 75, Appl
Sequence 473, App
Sequence 135, App
Sequence 135, App
Sequence 2128, Ap
Sequence 135, App
Sequence 1294, Ap
Sequence 135, App
Sequence 135, App
Sequence 1031, Ap
Sequence 135, App
Sequence 11, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 7, Appl
Sequence 6, Appl
Sequence 381, App
Sequence 9, Appl
Sequence 8, Appl
Sequence 10, Appl
Sequence 1, Appl

16 424.4 18.1 1539 12 US-10-332-447-39
17 416.4 17.8 2218 9 US-09-954-456-57
18 416.4 17.8 2218 9 US-09-954-456-596
19 416.4 17.8 2218 12 US-10-342-887-861
20 416.4 17.8 2218 15 US-10-325-027-307
21 376 16.0 1607 12 US-10-029-020-23
22 333.2 14.2 2135 12 US-10-332-447-47
23 330.2 14.1 2778 14 US-10-176-847-75
24 330.2 14.1 3366 14 US-10-101-510-473
25 328.6 14.0 2856 9 US-09-735-705-135
26 328.6 14.0 2856 9 US-09-850-716A-135
27 328.6 14.0 2856 9 US-09-880-107-2128
28 328.6 14.0 2856 9 US-09-897-778-135
29 328.6 14.0 2856 10 US-09-466-396A-135
30 328.6 14.0 2856 12 US-10-342-887-1294
31 328.6 14.0 2856 14 US-10-117-982-135
32 328.6 14.0 2856 15 US-10-325-027-1031
33 328.6 14.0 2856 15 US-10-313-986-135
34 328.6 14.0 2856 15 US-10-159-563-235
35 328.6 14.0 7777 14 US-10-067-449-11
36 325.4 13.9 2338 14 US-10-067-449-12
37 323.8 13.8 2338 14 US-10-067-449-13
38 322.2 13.8 2338 14 US-10-067-449-14
39 321.4 13.7 2613 9 US-09-778-927A-7
40 321.4 13.7 2683 9 US-09-778-927A-6
41 319.8 13.6 2349 9 US-09-822-849A-381
42 310.4 13.2 2860 9 US-09-778-927A-9
43 310.4 13.2 2919 9 US-09-778-927A-8
44 301.4 12.9 3046 9 US-09-778-927A-10
45 292.8 12.5 1918 9 US-09-822-863-1

ALIGNMENTS

RESULT 1
US-09-981-947A-1
; Sequence 1, Application US/09981947A
; Patent No. US20020164578A1
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/981,947A
; FILING DATE: 18-Oct-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2004, 14:27:39 ; Search time 819 Seconds
(without alignments)
10732.203 Million cell updates/sec

Title: US-09-981-947B-1
Perfect score: 2343
Sequence: 1 tcgaccacgcctcgccct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2475585 seqs, 1875730760 residues
Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2343	100.0	2343	9	US-09-981-947A-1
2	1651	70.5	2032	15	US-10-297-022-52
3	1644.6	70.2	1851	12	US-10-276-774-930
4	1399	59.7	1710	12	US-10-188-186-75
5	1317.8	56.2	2842	14	US-10-037-270-574
6	1317.8	56.2	2842	15	US-10-117-722-574
7	1232.8	52.6	1757	12	US-10-188-186-77
8	1229	52.5	1502	12	US-10-188-186-79
9	1187.6	50.7	1267	12	US-10-188-186-89
10	1141	48.7	2216	14	US-10-168-651-33
11	1133.8	48.4	1189	12	US-10-188-186-85
12	1130.6	48.3	1189	12	US-10-188-186-83
13	1130.6	48.3	1189	12	US-10-188-186-87
14	915	39.1	1087	12	US-10-188-186-81
15	430.8	18.4	472	10	US-09-918-995-30642

Search completed: April 8, 2004, 16:44:45
Job time : 196 secs

```

; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match      11.7%; Score 273.2; DB 2; Length 246240;
Best Local Similarity 86.3%; Pred. No. 1.7e-54;
Matches 302; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1983 AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCACGGTGGCTCAC 2042
Db 2949 AGAGATATAGATCTGTCTTTTATCTAAAAAGTCTTTATAAGCGCGGCGGTGGCTCAC 3008
QY 2043 GCCTGTAATCCAGACACTTTGGAGGCCGAGGTGGGTGAGTCTGTAGGTCAGGAGATTGA 2102
Db 3009 GCCTGTATATCCAGACACTTTGGAGGCCGAGGTGGGTGAGTCTGTAGGTCAGGAGATCGA 3068
QY 2103 GACCATCTGGCTTAACATGGTGAACACTCCCTCTCTACTAAAAATACAAAAAATTAGCTGA 2162
Db 3069 GACCATCTCCGGCTAAACAGCTGAACACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGG 3128
QY 2163 GCATGGTGGCGGCGGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCAGGAGAAATGCGCT 2222
Db 3129 GCGTAGTGGCGGCGGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCAGGAGAAATGCGCT 3188
QY 2223 GAACCCAGGAGCGGAGCTTTGCAGTGAGCGGAGATCACACCACTCCAGCTCCAGCTGGGT 2282
Db 3189 GAAACCGGAGGCGGAGCTTTGCAGTGAGCGGAGATCCCGCACTGCACCTCCAGCTGGGC 3248
QY 2283 GACAGAGCCAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAA 2332
Db 3249 GACAGAGCAAGACTCGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAA 3298

RESULT 15
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 587237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND AND TOWNSEND AND CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136

```

Matches 292; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2009 AAGCTTGAAGTCTAGGCTGTGACGGTGGCTACAGCTGTATCCAGCACTTTGGGAGG 2068
DB 320 AAAATGTTTCTGGGCGGGGGTGGCTACAGCTGTATCCAGCACTTTGGGAGG 261
QY 2069 CCGAGGTGGGTGGATCGTGGGTGAGGATTCAGCACTTCTGGCTAACTGTTGAAAC 2128
DB 260 CCGAGGTGGGTGGATCGTGGGTGAGGATTCAGCACTTCTGGCTAACTGTTGAAAC 201
QY 2129 TCCCTCTCTACTAAATAACAAAAATTAGCTGAGCATGGTGGGGGGCTGTAGTCCC 2188
DB 200 CCCGTCTCTACTAAATAACAAAAATTAGCTGAGCATGGTGGGGGGCTGTAGTCCC 141
QY 2189 AGCTACTTGGGAGGCTGAGGAGGATGGCTGAACCCAGGAGGCGGAGCTTCCAGTG 2248
DB 140 AGCTACTTGGGAGGCTGAGGAGGATGGCTGAACCCAGGAGGCGGAGCTTCCAGTG 81
QY 2249 AGCCGAGATCACACCACACTCCAGCTGGGTGACAGGCGAGCTCCGTTCTCAAAA 2308
DB 80 AGCCGAGATCCGCACTGCACTCCAGCTGGGCGAGAGCCAGCACTCCGTTCTCAAAA 21
QY 2309 AAAAAAATAAAAAAATAAAAAA 2328
DB 20 AAAAAAATAAAAAAATAAAAAA 1

RESULT 12
US-08-759-873-5/c
; Sequence 5, Application US/08759873
; Patent No. 5683885
; GENERAL INFORMATION:
; APPLICANT: Kieback, Dirk G.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AN INCREASED RISK
; TITLE OF INVENTION: OF BREAST OR OVARIAN CANCER
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MIION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,873
; FILING DATE: 12-APRIL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KIT, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6612
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
US-08-759-873-5

Query Match 11.7%; Score 275.2; DB 1; Length 320;
Best Local Similarity 91.2%; Pred. No. 5e-56;
Matches 292; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2009 AAGCTTGAAGTCTAGGCTGTGACGGTGGCTACAGCTGTATCCAGCACTTTGGGAGG 2068
DB 320 AAAATGTTTCTGGGCGGGGGTGGCTACAGCTGTATCCAGCACTTTGGGAGG 261
QY 2069 CCGAGGTGGGTGGATCGTGGGTGAGGATTCAGCACTTCTGGCTAACTGTTGAAAC 2128
DB 260 CCGAGGTGGGTGGATCGTGGGTGAGGATTCAGCACTTCTGGCTAACTGTTGAAAC 201
QY 2129 TCCCTCTCTACTAAATAACAAAAATTAGCTGAGCATGGTGGGGGGCTGTAGTCCC 2188
DB 200 CCCGTCTCTACTAAATAACAAAAATTAGCTGAGCATGGTGGGGGGCTGTAGTCCC 141
QY 2189 AGCTACTTGGGAGGCTGAGGAGGATGGCTGAACCCAGGAGGCGGAGCTTCCAGTG 2248
DB 140 AGCTACTTGGGAGGCTGAGGAGGATGGCTGAACCCAGGAGGCGGAGCTTCCAGTG 81
QY 2249 AGCCGAGATCACACCACACTCCAGCTGGGTGACAGGCGAGCTCCGTTCTCAAAA 2308
DB 80 AGCCGAGATCCGCACTGCACTCCAGCTGGGCGAGAGCCAGCACTCCGTTCTCAAAA 21
QY 2309 AAAAAAATAAAAAAATAAAAAA 2328
DB 20 AAAAAAATAAAAAAATAAAAAA 1

RESULT 13
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240

Query Match 11.8%; Score 276.6; DB 4; Length 12619;
Best Local Similarity 89.7%; Pred. No. 8.9e-56;
Matches 297; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 2003 TTCTAAAGCTTGAAGTCTAGGCTGTGACGGTGGCTCAGCCTGTAAATCCAGCACTTT 2062
Db 10625 TTTTAAAAATCTGCAGCATAGGCGCGGCGGTGCTGACGCTGTAAATCCAGCACTTT 10566
Qy 2063 GGGAGCCGAGGTTGGTGGTCTGTAGGTCAGAGATTGAGACATCCTGGCTAACATGG 2122
Db 10565 GGGAGCCGAGGCGCGGCGGATCAGAGGTGAGAGATCAAGCATCTCTGGCTAACATGG 10506
Qy 2123 TGAACCTCCCTCTCTACTAAAAATACAAAAATAGCTGAGCATGTGGCGGCGGCTGT 2182
Db 10505 TGAACCCCGCTCTCTACTAAAAATACAAAAATAGCTGAGCATGTGGCGGCGGCTGT 10446
Qy 2183 AGTCCCACTTCTGAGGCTGAGCAGGAGATGCGCTGAAACCCAGGAGGCGGCTTT 2242
Db 10445 AGTCCCACTTCTGAGGCTGAGCAGGAGATGCGCTGAAACCCAGGAGGCGGCTTT 10386
Qy 2243 GCAGTGGCCGAGATCACACCACTCCAGCTGGGTGACAGGCGGAGACTCCGCTCT 2302
Db 10385 GCAGTGGCCGAGATTGCGCCACTGCACTCCAGCTGGGCGGAGCAAGACTCCGCTCT 10326
Qy 2303 CAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2333
Db 10325 CAAAAAATAAAAAAAAAAAAAAAAAAAAAA 10295

RESULT 11
US-08-629-939-5/c
; Sequence 5, Application US/08629939
; Patent No. 5645995
; GENERAL INFORMATION:
; APPLICANT: Kieback, Dirk G.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AN INCREASED
; RISK OF BREAST OR OVARIAN CANCER
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/629,939
; FILING DATE: 12-APRIL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KIT, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: A-6612
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
US-08-629-939-5

Query Match 11.7%; Score 275.2; DB 1; Length 320;
Best Local Similarity 91.2%; Pred. No. 5e-56;

Qy 1018 CAGGTGGTCAACGATGTTACCATGGCTGCTACAGCTCTGTGGCTCAATGCAATT 1077
Db 990 CCCATCCTCATCGCTGTGTGCTGAGCTGTCCAGCAGCTGTCTGGCATCAACGCTGTC 1049
Qy 1078 TGGTCTTACCAACAGAGATCTTTGAAAAAGCTGGGATCCCTCCGCGAAAGATCCCATAC 1137
Db 1050 TTCTATTACTCCAGAGCATCTTCAGAAAGCGGGGGTGCAGACCCCTGTG-----TAT 1103
Qy 1138 GTCACTTGTAGTACAGGGGGCATCAGAGCTTTGGTGGCTGCTTCTCTGTTTGGTCAAT 1197
Db 1104 GCCACATTGGTCTCGGTATCTGCAACAGGCTTCACTGTGCTGCTGTTTGGTGTG 1163
Qy 1198 GAGCAGCTGGAGAGAGACCCCTCTCTCAATGTGTGGCTTTGGGCTCATGGGCTCTTCTTT 1257
Db 1164 GAGCAGCAGGCGCGGAGACCCCTGCACCTCATAGGCTCTGCTGGCATGGCGGTTGTGCC 1223
Qy 1258 GGGAGCCCTACATACGCTGACCCCTGAGGACGAGCCGCTGGTGGCTTCTCTCTGAGT 1317
Db 1224 ATACTCATGACCATCGCTAGCTGCTGGAGCAGCTACCCCTGGATGCTCTATCTGAGC 1283
Qy 1318 ATCGTGGGATCTGCGCATATCGCTCTTTCTGAGTGGCGGAGGATCCCGTTC 1377
Db 1284 ATCGTGGCATCTTTGGCTTTTGGCTTTCTTTGAAGTGGTCTCTGGCCCATCCATGG 1343
Qy 1378 ATCTTGACTGTGAGTCTTCCAGCAATCTCAGCGCGGCTGCTTCCATTCATTCAGGAC 1437
Db 1344 TTATCGTGGTGAATCTTTCAGCAGGCTCCAGCTCCAGCTGCCATTCGCGTTGCGAGC 1403
Qy 1438 ACCGTCACTGCTCTCCAACTTCTGCTGTGGCTCCCTTCCATTCATTCAGAAAAAGT 1497
Db 1404 TTCTCCAACTGAGCTCAAAATTCATTTGGGCTGCTTCCAGTGTGGAGCAACTG 1463
Qy 1498 CTGGACACTACTGTTTCTAGTCTTTGCTACAAATTTGATCAGAGTGTCTATCTACCTG 1557
Db 1464 TGTGTCCTTCTTCTATCTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
Qy 1558 TATTTTGTGCTGCTGAGACCAAAACAGAACTATGAGAAATC 1602
Db 1524 TACTTCAAAGTCTCTGAGACTAAAGCCGAGCTTCGATGAGATC 1568

RESULT 10
US-09-616-289-49/c
; Sequence 49, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 12619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-49

Matches 725; Conservative 0; Mismatches 634; Indels 6; Gaps 1;
Qy 238 CTCCTGTTGGCTCCCTCGGGGGCGCTTGGGCTCCTCTTCTCTACGGCTACACCTG 297
Db 210 CGCTTCATGCTGCTGAGGAGAGAGTCTTGGCTCCCTGAGTTGGCTTACACACT 269
Qy 298 TCGGTGTTGAATGCCCGCCACCCCGTACATCAAGCCCTTTTACATGAGTCAATGGGAAGA 357
Db 270 GGAGTCATCAATGCCCGCCACAGAGGTGATCGAGAGTTCTCAACACAGACATGGGTCCAC 329
Qy 358 AGGCATGGAGCTCCCAATAGACCCAGACACTCTGACTCTCTGCTCTGGTCTGTGACTGTCTC 417
Db 330 CGCTATGGGAGAGACATCTGCCACACAGCTCACACGCTCTGGTCCCTCTCAGTGGCC 389
Qy 418 ATATGCCCATCGGTGGAGCTTGGGGAGCTTAATTTGTAAGATGATTTGGAAGTTCTT 477
Db 390 ATCTTTTCTGTTGGGGGATGATGGCTCTCTCTGTTGGGCTTTTCGTAAACCGCTT 449
Qy 478 GGGAGAGACACTTTTGTGTCGCAATATGGGTTTGGCAATTTCTGCTGCAATGCTGATG 537
Db 450 GCGCGCGGAATTCATGCTGATGATGAACCTGCTGGCCCTTCGTGTCCGCGTGTCTATG 509
Qy 538 GCTGCTGCTCCAGCAGGAGCTTTGAAATGCTCATTTGTCGGAGCTTCTCATATGGGC 597
Db 510 GGCTTCTCGAACTGGGCAAGTCCCTTTGAGATGCTGATCCTGGGCGCTTCTATCATCGGT 569
Qy 598 ATAGATGGAGGGTCCCTCAGTGTCTCCCATGTACTCTCAGTGAGATCTCACCCAG 657
Db 570 GTGACTGGGCTGACACAGCTTCGTGCCATGATGTGGTGAAGTGTACCCACA 629
Qy 658 GAGATCCGTGCTCTGCGGCGAGTGTACTGCCATCTTTATCTGATGGCGTGTACT 717
Db 630 GCCTTTCTGTTGGGCGCTGGCCCTGACAGCTGGGCTGGCATGCTGTGGCATCTCATC 689
Qy 718 GGGAGCTTCTGGGCTCCCGAGCTGTGGGAAGGAGAGTACTGGCCATACCTGTTT 777
Db 690 GCGCAGGTTCGGCTGTGACTCCATCATGGGCAAGAGCTGTGGCCCTGTGCTG 749
Qy 778 GAGTGTATGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
Db 750 AGCATCATCTTATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
Qy 838 CCACGCTACTGCTCTTGGAGAAGCACACAGGAGAGAGCTGTGAAAGCTTCCAAACG 897
Db 810 CCGCGTCTCTGCTCATCAACGCAACGAGAGAGACCGGGCCAAAGTGTGCTAAAGAG 869
Qy 898 TTCTTGGTAAAGCAGAGCTTTCCCAAGAGGTAGAGAGTCTCTGGCTGAGAGCCACGTG 957
Db 870 CTGGCGGAGACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGAGAGAGTGGCAG 929
Qy 958 CAGAGAGCATCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017
Db 930 ATGATCGGGGAGAAGAGTCAACATCTCTGGAGCTGTTCCGCTCCCGGCTTACCGCCAG 989
Qy 1018 CAGGTGCTACCGTGTATTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
Db 990 CCAATCCTCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1049
Qy 1078 TGTCTTATACCAACAGCAGCTTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATAC 1137
Db 1050 TTCTATTATCCACAGCATCTTCGAGAGGGGGGTGCAGACCTGTG-----TAT 1103
Qy 1138 GTCACTTGAATCAGAGGGGCAATGAGACTTTGGCTGCGCTTCTCTGTTTGTGCTATT 1197
Db 1104 GCAACATTGGTCCGGTATCGTCAACACGGCTTCACTGTGCTGCTGCTGCTGCTGCTG 1163
Qy 1198 GAGCCTGGGAGGAGACCTTCTCATGTTGGCTTTGGGCTCATGGGCTCTTCTT 1257
Db 1164 GAGCAGAGGCGGGGAGCCCTGACCTCATAGGCTCTGCTGGCATGGCGGTTGTGCC 1223
Qy 1258 GGGACCTCACCATCAGCTGACCTGAGGAGACCGCCCTGGGTCCCTTACCTGAGT 1317
Db 1224 ATACTCATGACCATCGCGCTAGACTGTCTGGAGAGCTACCTCGTGTATCTGCTGCTG 1283

Qy 1318 ATCTGGGCACTTCGGGCATCATCGCTCTTTCTGCTGAGTGGGCCAGGTGGCATCCCGTTC 1377
Db 1284 ATCTGGGCCATCTTTTGGCTTTTGGCTCTTTTGAAGTGGGTCTGGCCCCATCCCATGG 1343
Qy 1378 ATCTTGACTGTGAGTTCTTCCAGCAATCTCAGCGGCGGCTGCTTTCATCATTTGCAGGC 1437
Db 1344 TCTATCGTGGCTGAATCTTTCAGCAGGTCCAGCTCCAGCTGCTGCTGCTGCTGCTGCTG 1403
Qy 1438 ACCGTCAACTGGCTCTCCAACTTTTGTGTTGGGCTCTCTTCCCATTCATTTCAGAAAAGT 1497
Db 1404 TTCTTCAACTGGACCTCAATTTTCATTTGTGGCATGTCTTCCAGTATGTGGACCACTG 1463
Qy 1498 CTGACACCTACTGTTTCTTCTAGTCTTTGTCTACAAATTTGTATACAGGTGTCTACTG 1557
Db 1464 TGTGCTCCCTACGCTTTCATCATCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
Qy 1558 TATTTTGTGCTGCTGAGACCAAAAACAGAACCTTATGCAAAATC 1602
Db 1524 TACTTCAAAGTTCCTGAGACTAAAGCGCGACCTTCGATGAGATC 1568

RESULT 8
US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-135

Query Match 14.08; Score 328.6; DB 4; Length 2856;
Best Local Similarity 53.14; Pred. No. 2.5e-68;
Matches 725; Conservative 0; Mismatches 634; Indels 6; Gaps 1;
Qy 238 CTCCTGTTGGCTCCCTCGGGGGCGCTTGGGCTCCTCTTCTCTACGGCTACACCTG 297
Db 210 CGCTTCATGCTGCTGAGGAGAGAGTCTTGGCTCCCTGCGATTGGCTACCACT 269
Qy 298 TCGGTGTTGAATGCCCGCCACCCCGTACATCAAGCCCTTTTACATGAGTCAATGGGAAGA 357
Db 270 GGAGTCATCAATGCCCGCCACAGAGGTGATCGAGAGTTCTACAACACAGACATGGGTCCAC 329
Qy 358 AGGCATGGAGCTCCCAATAGACCCAGACACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCT 417
Db 330 CGTATGGGAGAGCATCTTGGCCGCCACCGCTCACACGCTTCCCTGCTGCTGCTGCTGCTGCT 389
Qy 418 ATATTCGCCATCGGTGAGCTTGGGAGCTTAAATGTGAAGATGATTGGAAGGTTCTT 477
Db 390 ATCTTTTCTGTTGGGGGATGATGGCTCTCTCTGTTGGGCTTTTCGTAAACCGCTT 449
Qy 478 GGGAGGAGACACTTTTGTCTGGCAATAATGGGTTTGGCAATTTCTGCTGCAATGCTGATG 537
Db 450 GCGCGCGGAATTCATGCTGATGATGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509

538 GCCTGCTGCTCAGGACGAGCCTTTGAAATGCTCATTTGTGGGAGCCTTCATCATGCGC 597
510 GGCCTTCTCGAAATGGGCAAGTCCCTTTGAGATGCTGATCCTGGGCGCCTTCATCATCGGT 569
598 ATAGATGGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAG 657
570 GTGTACTGCGGCGCTGACACAGCCTTCGTGCCCATGTATGTGGGTGAGTGTCAACCCACA 629
658 GAGATCCGTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCATTTGGCGTGTTCACCT 717
630 GCCTTTCGTGGGCGCTCTGGGCAACCTGCACAGCTGGGCATCGTCTGGGCATCCTCATC 689
718 GGGCAGCTTCTGGGCGCTCCCGAGCTGCTGGGAAAGGAGATGACTTGGCCATACCTGTTTT 777
690 GCCCAGGTCTTGGCGTGGACTCCATCATGGGCAACAAGAACCTGTGGCCCTCTGCTGCTG 749
778 GGAGTGATTGTGGTCCCTCGCGTGTGCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGC 837
750 AGCATCATCTTATCCCGCGCTGCTGCGAGTGTCATCGTGTGCCCTCTCTGCCCGAGAGT 809
838 CCACGCTACTGCTCTTGGGAAGACACAACGAGGCAAGAGCTGTGAAAGCCTTCCAAACG 897
810 CCCCGCTTCTGCTCATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCTAAGAAG 869
898 TTCTTTGGGTAAAGCAGACGCTTTTCCAAAGAGGTAGAGAGGTCTCTGCTGAGAGCCACGCTG 957
870 CTGGCGGGACAGCTGACGTGACCCATGACCTGTCAGGAGATGAAGGAAGAGAGTCTGGCAG 929
958 CAGAGAGCATCCGCGCTGTGTGCTGCTGAGAGCTGCTGAGAGCTCCCTACGTCCCGCTGG 1017
930 ATGATGCGGGAGAAGAGGTACCATCTCTGGAGCTGTTCGCGCTCCCGCGCTACCGCCAG 989
1018 CAGGTGTCACCGTGATTGTCACCATGGCGCTGTACAGCTCTGTGCGCTCAATGCAATT 1077
990 CCATCCTCATCGCTGTGTGTGCTGAGCTGTCCAGAGCTGTCTGCATCAAGCTGTCT 1049
1078 TGTCTTATACCAACAGCATCTTTGAAAAGCTGGGATCCCTCCGGCAAAAGATCCCAATC 1137
1050 TTCTATTATCCACAGACATCTTCGAGAGCGGGGGTGCAGCAGCCTGTG-----TAT 1103
1138 GTCACCTTGAGTACAGGGGGCATTCGAGACTTTGGCTGCGGTCTTCTCTGTTTGGTCAATT 1197
1104 GCCACCATTTGGCTCCGGTATCGTCAACACGGCCTTCACTGTCTGTCTGCTGTTTGTGGTG 1163
1198 GAGACCTGGAGCGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCGCTCTTCTTT 1257
1164 GAGCGAGCAGCGCGCGGACCCCTGCACCTCATAGGCGCTCGCTGGCATGGCGGGTGTGACC 1223
1258 GGGACCTCACCATCAGCTGACCTTCGAGGACCAAGCCCGCTGGGTCCCTACTCTGAGT 1317
1224 ATACTCATGACCATCGCTAGACCTGCTGGAGAGCTACCTTGAGTGTCTTATCTGAGC 1283
1318 ATCGTGGGCAATCTGGCCATCATCGCTCTTTCTGCAGTGGGCGAGGTGCATCCCGCTTC 1377
1284 ATCGTGGCCATCTTTGGCTTTTGGGCTTCTTTGAAGTGGGTCTTGCCCGCCATCCCATGG 1343
1378 ATCTTGACTGGTGAGTTCTTCCAGCAATCTCAGCGGCGGGCTGCTTCATCATTTGCGAGGC 1437
1344 TTCACTGCTGGCTGAATCTTTCAGCCAGGTTCACGTCCAGCTGCCATTTGCCGTGCGAGGC 1403
1438 ACCGTCAACTGGCTTCCAACTTTGTGTGGGCTCTCTTCCCATTCATTCAGAAAAGT 1497
1404 TTCTCAACTGGACCTCAAAATTTCAATTTGGGCATGTGCTTCCAGTATGTGGAGCAACTG 1463
1498 CTGGACACCTACTGTTTCTAGTCTTTGTACAAATTTGATTCACAGGTGCTATCTACTG 1557
1464 TGTGTTCCCTACTGCTTTCATCTTCTTCACTGTGCTCTGTGTTCTGTCTTCTCATCTTCACC 1523
1558 TATTTTGTGTGCTGCTGAGACCAAAAAACAGAACCTATGCAGAAATC 1602
1524 TACTTCAAAGTTCCTGAGACTAAAGCCGAGACCTTCGATGAGATC 1568

Db	1261	ACCCTCACCATCAGCTGACCCCTGCAGGACCAAGCCCCCTGGGTCCCTACCTGAGTATC
Qy	1321	GTGGGCAATCTGGCCATCATCGGCTCTTTCTGCAAGTGGCGCAGGTGGCATCCCGTTTCATC
Db	1321	GTGGGCAATCTGGCCATCATCGGCTCTTTCTGCAAGTGGCGCAGGTGGCATCCCGTTTCATC
Qy	1381	TTGACTGGTGAGTTCTTCCAGCAATCTCAGCGCCGGCTGCCCTTCATCATTTGACAGGCAAC
Db	1381	TTGACTGGTGAGTTCTTCCAGCAATCTCAGCGCCGGCTGCCCTTCATCATTTGACAGGCAAC
Qy	1441	GTCAACTGGCTCTCGAACTTTTGGCTCTTGGCTCTTCCCATTCATTCAGAAAAGTCTG
Db	1441	GTCAACTGGCTCTCGAACTTTTGGCTCTTGGCTCTTCCCATTCATTCAGAAAAGTCTG
Qy	1501	GACACCTACTGTTTCTTAGTCTTTGTACAAATTTGTATCACAGGTGCTATCTACCTGTAT
Db	1501	GACACCTACTGTTTCTTAGTCTTTGTACAAATTTGTATCACAGGTGCTATCTACCTGTAT
Qy	1561	TTTGTGCTGCTGAGACCAAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCAAA
Db	1561	TTTGTGCTGCTGAGACCAAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCAAA
Qy	1621	AGGAACAAAGCATACCCACAGAAAGAAAATCGACTCAGCTGTCACTGATCTCTCTGCT
Db	1621	AGGAACAAAGCATACCCACAGAAAGAAAATCGACTCAGCTGTCACTGATCTCTCTGCT
Qy	1681	TCCTTCCTTTCTACTCTCGAATACAGCCCTGGATCAAGCTCGCGCCACCAACACCGCC
Db	1681	TCCTTCCTTTCTACTCTCGAATACAGCCCTGGATCAAGCTCGCGCCACCAACACCGCC
Qy	1741	ACCAAAAAGAACACCCCATTTGTAACCGGTCAATGTTGTTATTTCTTCAACCTGGAAATGACCT
Db	1741	ACCAAAAAGAACACCCCATTTGTAACCGGTCAATGTTGTTATTTCTTCAACCTGGAAATGACCT
Qy	1801	TCCCTATCTTCTCTCGGAAACACCAAGTCATGATGTACAGAAAGCTTTGGATTT
Db	1801	TCCCTATCTTCTCTCGGAAACACCAAGTCATGATGTACAGAAAGCTTTGGATTT
Qy	1861	TGGAGACATGGTTTGAATTCAGTCAATCTTTTATTCAGCAAAATATTTAACAAAGT
Db	1861	TGGAGACATGGTTTGAATTCAGTCAATCTTTTATTCAGCAAAATATTTAACAAAGT
Qy	1921	ACTGACATGCCATATGTTGTTTATCCACTGTTTATACAACTGGGAGGAGAGAGAG
Db	1921	ACTGACATGCCATATGTTGTTTATCCCACTGGTTATACAACTGGGAGGAGAGAGAG
Qy	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTTGAAAGCTGTAGGCTGTGCACGGTGGCTC
Db	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTTGAAAGCTGTAGGCTGTGCACGGTGGCTC
Qy	2041	ACGCTGTAAATCCAGACACTTTGGAGGCGCAGGTGGGTGGATCGTGAGGTTCAGAGATT
Db	2041	ACGCTGTAAATCCAGACACTTTGGAGGCGCAGGTGGGTGGATCGTGAGGTTCAGAGATT
Qy	2101	GAGACCATCTCTGGCTAAACATGGTGAAACTCCCTCTCTACTAAAAATACAAAAAATTAGCT
Db	2101	GAGACCATCTCTGGCTAAACATGGTGAAACTCCCTCTCTACTAAAAATACAAAAAATTAGCT
Qy	2161	GAGCATGGTGGCGGCGCTGTAGTCCAGTACTTTGGAGGCTGAGCAGGAGAAATGCG
Db	2161	GAGCATGGTGGCGGCGCTGTAGTCCAGTACTTTGGAGGCTGAGCAGGAGAAATGCG
Qy	2221	GTGAACCCAGAGCGGAGCTTCAGGTGAGCGGAGATCACACCACTCCAGCTCTGG
Db	2221	GTGAACCCAGAGCGGAGCTTCAGGTGAGCGGAGATCACACCACTCCAGCTCTGG
Qy	2281	GTGACAGCCAGACTCCGCTCTCAAAAAAATAAAAAAAAAAAAAAAAAAGGGCGGC
Db	2281	GTGACAGCCAGACTCCGCTCTCAAAAAAATAAAAAAAAAAAAAAAAAAGGGCGGC
Qy	2341	CGC 2343

2341 CGC 2343

```

Db      2281 GTGACAGACGAGCTCCGCTCTCAAAAAAAAAAAAAAAAAAGGCGC 2343
QY      2341 CGC 2343
Db      2341 CGC 2343

RESULT 3
US-09-610-417-1
; Sequence 1, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkietohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2343 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 73...1761
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-610-417-1

Query Match          100.0%; Score 2343; DB 4; Length 2343;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Db 1 TCAGCCACCGCTCCGGCTTGGCAGAGTCTGGGTCCCTGGACTGAGCCATCAGCTGGG 60
Qy 61 TCACAGAGCCCATGAGCAAGGAAACAAATAAGAAATTCACAGGAACCTAGTTCCC 120
Db 61 TCACAGAGCCCATGAGCAAGGAAACAAATAAGAAATTCACAGGAACCTAGTTCCC 120
Qy 121 CTCACAGATGACACACAGCCCGGCTCCAGGCGCAGGAGGAGGAGCTCTGGAGTGT 180
Db 121 CTCACAGATGACACACAGCCCGGCTCCAGGCGCAGGAGGAGGAGCTCTGGAGTGT 180
Qy 181 GACACCTCAGGAGTGGGTGGCAGGTGGAAGGAGAAAGAAAGAACTGGTCTGGCTC 240
Db 181 GACACCTCAGGAGTGGGTGGCAGGTGGAAGGAGAAAGAAAGAACTGGTCTGGCTC 240
Qy 241 CTGGTGGCTCCCTCCGGGGGCTTGGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 241 CTGGTGGCTCCCTCCGGGGGCTTGGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Qy 301 GTGGTGAATGCCCCACCCGCTACATCAGGCGCTTTTCAATGAGTGTGGAAGAGG 360
Db 301 GTGGTGAATGCCCCACCCGCTACATCAGGCGCTTTTCAATGAGTGTGGAAGAGG 360
Qy 361 CATGAGCGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCATA 420
Db 361 CATGAGCGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCATA 420
Qy 421 TTGCGCATCGGTGGAGCTTGTGGGAGCTTAAATGTGAAGATGATGGAAGAGTCTTGGG 480
Db 421 TTGCGCATCGGTGGAGCTTGTGGGAGCTTAAATGTGAAGATGATGGAAGAGTCTTGGG 480
Qy 481 AGGAAGCACACTTGTGGCGCAATATGGTGTGCAATTTCTGCTGCTCTCTCTCTCTCTCT 540
Db 481 AGGAAGCACACTTGTGGCGCAATATGGTGTGCAATTTCTGCTGCTCTCTCTCTCTCTCT 540
Qy 541 TGCTGCTCCAGGAGGAGGCTTTGAAATGCTCATTTGTGGGAGCTTTCATCATGGGATA 600
Db 541 TGCTGCTCCAGGAGGAGGCTTTGAAATGCTCATTTGTGGGAGCTTTCATCATGGGATA 600
Qy 601 GATGAGAGGCTCGCCCTCAGTGTGCTCCCAAGTACCTCAGTGAATCTACCCAGGAG 660
Db 601 GATGAGAGGCTCGCCCTCAGTGTGCTCCCAAGTACCTCAGTGAATCTACCCAGGAG 660
Qy 661 ATCCGTGGCTCTCTGGGAGGAGTACTGCCATCTTTATCTCANTGGGCTTCACTGGG 720
Db 661 ATCCGTGGCTCTCTGGGAGGAGTACTGCCATCTTTATCTCANTGGGCTTCACTGGG 720
Qy 721 CAGCTTCTGGGCTCGCCGAGCTGCTGGGAAAGGAGTACTGCGCCATACCTGTTTGA 780
Db 721 CAGCTTCTGGGCTCGCCGAGCTGCTGGGAAAGGAGTACTGCGCCATACCTGTTTGA 780
Qy 781 GTGATGTGGTCCCTCGGCTTCCAGCTGTGAGCTTCCCTTCTCCGCGAGAGCCCA 840
Db 781 GTGATGTGGTCCCTCGGCTTCCAGCTGTGAGCTTCCCTTCTCCGCGAGAGCCCA 840
Qy 841 CGCTACCTCTCTTGGAGAGCACAAACAGAGCAGAGCTGTGAAGCTTCCAAACGCTC 900
Db 841 CGCTACCTCTCTTGGAGAGCACAAACAGAGCAGAGCTGTGAAGCTTCCAAACGCTC 900
Qy 901 TTGGTAAAGCAGAGCTTTCCCAAGAGGTAGAGAGGCTCTGGCTGAGAGCAGCTGCAG 960
Db 901 TTGGTAAAGCAGAGCTTTCCCAAGAGGTAGAGAGGCTCTGGCTGAGAGCAGCTGCAG 960
Qy 961 AGAGCATCCGCTGGTGTCTGTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCAG 1020
Db 961 AGAGCATCCGCTGGTGTCTGTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCAG 1020
Qy 1021 GTGGTCAACGCTGATGTACCAATGAGCTTACAGCTCTGTGGCTCAATGCAATTTGG 1080
Db 1021 GTGGTCAACGCTGATGTACCAATGAGCTTACAGCTCTGTGGCTCAATGCAATTTGG 1080
Qy 1081 TTCTATACCAACAGCATCTTTGAAAGCTGGGATCCCTCCGGCAAGAGTCCCATACGTC 1140

Db 1081 TTCTATACCAACAGCATCTTTGGAAGAGCTGGATCCCTCCGGCAAGAGTCCCATACGTC 1140
Qy 1141 ACCTTGTAGTACAGGGGATCGAGACTTTTGGCTGCCGCTTCTCTGTGTTGGTCAATTGAG 1200
Db 1141 ACCTTGTAGTACAGGGGATCGAGACTTTTGGCTGCCGCTTCTCTGTGTTGGTCAATTGAG 1200
Qy 1201 CACCTGGGACGAGAGCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGGG 1260
Db 1201 CACCTGGGACGAGAGCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGGG 1260
Qy 1261 ACCCTCACCATCAGCCTGACCTTGACAGCACCCCTCCCTGGTCCCTCTACTCTGAGTATC 1320
Db 1261 ACCCTCACCATCAGCCTGACCTTGACAGCACCCCTCCCTGGTCCCTCTACTCTGAGTATC 1320
Qy 1321 GTGGGATTTGGGCTCATCTGCTCTTTCTGAGTGGGCGCAGGTGGATCCCTTCTATC 1380
Db 1321 GTGGGATTTGGGCTCATCTGCTCTTTCTGAGTGGGCGCAGGTGGATCCCTTCTATC 1380
Qy 1381 TTGACTGTGTGAGTCTTCCAGCAATCTCAGCGCGGCTGCTCTCATCTTTCAGGAGCACC 1440
Db 1381 TTGACTGTGTGAGTCTTCCAGCAATCTCAGCGCGGCTGCTCTCATCTTTCAGGAGCACC 1440
Qy 1441 GTCAACTGGCTCTCCAACTTTTGGGCTCTCTTCCCATTCATTTCAGAAAAGTCTG 1500
Db 1441 GTCAACTGGCTCTCCAACTTTTGGGCTCTCTTCCCATTCATTTCAGAAAAGTCTG 1500
Qy 1501 GACACTACTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
Db 1501 GACACTACTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
Qy 1561 TTTTGTGCTGCTGAGACCAAAACAGAACCTATCGAGAAATCAGCAGGAGATTTTCCAAA 1620
Db 1561 TTTTGTGCTGCTGAGACCAAAACAGAACCTATCGAGAAATCAGCAGGAGATTTTCCAAA 1620
Qy 1621 AGGAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTACTGTCTCTGCT 1680
Db 1621 AGGAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTACTGTCTCTGCT 1680
Qy 1681 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
Db 1681 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
Qy 1741 ACCAAAAAGAACACCCATCTGAAAGCTCATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
Db 1741 ACCAAAAAGAACACCCATCTGAAAGCTCATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
Qy 1801 TCCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
Db 1801 TCCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
Qy 1861 TGGAGACATGGGTTTGAATTCAGTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
Db 1861 TGGAGACATGGGTTTGAATTCAGTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
Qy 1921 ACTGACATGCTCCCATATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Db 1921 ACTGACATGCTCCCATATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Qy 1981 AGAGAGAGAGAGAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2040
Db 1981 AGAGAGAGAGAGAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2040
Qy 2041 AGCCTGTAAATCCAGACACTTTGGAGGCGGAGTGGTGGATCGTGTAGGTCTAGGAGATT 2100
Db 2041 AGCCTGTAAATCCAGACACTTTGGAGGCGGAGTGGTGGATCGTGTAGGTCTAGGAGATT 2100
Qy 2101 GAGACATCTCTGGCTTAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2160
Db 2101 GAGACATCTCTGGCTTAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2160
Qy 2161 GAGCATGTGTGGGCGGCTGTAGTCCAGCTACTTGTGGAGGCTGTAGGAGGAGATGGC 2220
Db 2161 GAGCATGTGTGGGCGGCTGTAGTCCAGCTACTTGTGGAGGCTGTAGGAGGAGATGGC 2220

OM nucleic - nucleic search, using sw model

Run on: April 8, 2004, 11:05:50 ; Search time 176 seconds
(without alignments)
7387.784 Million cell updates/sec

Title: US-09-981-947B-1
Perfect score: 2343
Sequence: 1 tcgaccacgcgtcgccct.....aaaaaaagggcgccgc 2343

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2343	100.0	2343	2	US-09-031-392-1
2	2343	100.0	2343	3	US-09-299-549-1
3	2343	100.0	2343	4	US-09-610-417-1
4	1317.8	56.2	2842	4	US-09-620-312D-574
5	328.6	14.0	2856	4	US-09-643-597-135
6	328.6	14.0	2856	4	US-09-480-884A-135
7	328.6	14.0	2856	4	US-09-542-615A-135
8	328.6	14.0	2856	4	US-09-606-421B-135
9	328.6	14.0	2856	4	US-09-221-107-135
C 10	276.6	11.8	12619	4	US-09-616-289-49
C 11	275.2	11.7	320	1	US-09-629-939-5
C 12	275.2	11.7	320	1	US-08-759-873-5
13	273.2	11.7	246240	2	US-08-724-394A-20
14	273.2	11.7	246240	2	US-08-724-394A-21
15	273.2	11.7	246240	2	US-08-724-394A-22
16	270.4	11.5	193303	4	US-09-497-855A-37
17	270.4	11.5	193303	4	US-09-497-855A-44
C 18	269	11.5	8220	4	US-09-797-908-3
C 19	268.6	11.5	14636	3	US-09-173-914-6
20	268.4	11.5	3742	1	US-08-694-915-5
C 21	266.4	11.4	116592	4	US-09-818-512-3
22	266	11.4	62804	4	US-09-800-960-3
23	266	11.4	62804	4	US-10-096-960-3
24	264.4	11.3	2592	4	US-09-591-025-8
25	264.4	11.3	2592	4	US-08-894-927B-8
26	263	11.2	22481	4	US-08-367-841A-43
27	263	11.2	22481	5	PCT-US95-07201-43

28	263	11.2	22484	4	US-09-875-223-2	Sequence 2, Appli
29	263	11.2	22484	4	US-09-875-114-2	Sequence 2, Appli
C 30	262	11.2	38564	4	US-09-734-673-3	Sequence 3, Appli
C 31	262	11.2	319608	4	US-09-539-333D-1	Sequence 1, Appli
C 32	262	11.2	319608	4	US-09-679-409-1	Sequence 1, Appli
33	261.8	11.2	14796	3	US-08-975-080-35	Sequence 35, Appli
34	261.8	11.2	14796	3	US-09-630-706-10	Sequence 3, Appli
35	261.8	11.2	14796	3	US-09-496-694B-3	Sequence 3, Appli
36	261.4	11.2	856	4	US-09-288-143-47	Sequence 47, Appli
37	261.4	11.2	14581	4	US-08-520-373D-4	Sequence 4, Appli
C 38	260.6	11.1	588	3	US-09-385-982-128	Sequence 128, App
C 39	260	11.1	84495	4	US-09-797-906-3	Sequence 3, Appli
40	258.8	11.0	10825	3	US-08-652-265-1	Sequence 1, Appli
41	258.8	11.0	10825	3	US-08-652-265-3	Sequence 3, Appli
42	258.8	11.0	10825	3	US-08-652-265-5	Sequence 5, Appli
43	258.8	11.0	10825	3	US-08-652-265-7	Sequence 7, Appli
44	258.8	11.0	10825	3	US-08-834-497A-1	Sequence 1, Appli
45	258.8	11.0	10825	3	US-08-834-497A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-031-392-1
; Sequence 1, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 25-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2343 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 73...1761
US-09-031-392-1

Query Match 100.0%; Score 2343; DB 2; Length 2343;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCGACCCACGCGTCCGCGCTTCGAGTCTCGGGTCCCTGAGTCATGACCATGCTGGG 60

QY 397 CTCTGGTCTGTGACGTGTCTCCATATTCGCCATCGGTGGACCTTGTGGGACGTTAATTGTG 456
Db |||||
QY 397 CTCTGGTCTGTGACGTGTCTCCATATTCGCCATCGGTGGACCTTGTGGGACGTTAATTGTG 396
Db |||||
QY 457 AAGATGATGGAAGGTTCTTGGGAGGAGCACACTTTCCTGGCCATATATGGGTTTGC 516
Db |||||
QY 397 AAGATGATGGAAGGTTCTTGGGAGGAGCACACTTTCCTGGCCATATATGGGTTTGC 456
Db |||||
QY 517 ATTTCTGTCTGATTCCTGATGCTCTCGCTCCAGGAGGAGCTTGTGAATGCTCAT 576
Db |||||
QY 457 ATTTCTGTCTGATTCCTGATGCTCTCGCTCCAGGAGGAGCTTGTGAATGCTCAT 516
Db |||||
QY 577 GTGGACGCTTCATCATGGGCATAGATGGAGCGCTCGCCCTCAGTGTGCTCCCCATGTAC 636
Db |||||
QY 517 GTGGACGCTTCATCATGGGCATAGATGGAGCGCTCGCCCTCAGTGTGCTCCCCATGTAC 576
Db |||||
QY 637 CTCAGTGATCTACCCAGAGATCCGTGGCTCTCTGGGAGGAGCTGCGCATCTTT 696
Db |||||
QY 577 CTCAGTGATCTACCCAGAGATCCGTGGCTCTCTGGGAGGAGCTGCGCATCTTT 636
Db |||||
QY 697 ATCTGCATGGGCTTCACTGGGAGCTTCTGGGCTGCGGAGCTCTGGGAAAGGAG 756
Db |||||
QY 637 ATCTGCATGGGCTTCACTGGGAGCTTCTGGGCTGCGGAGCTCTGGGAAAGGAG 696
Db |||||
QY 757 AGTACTGGCCATACCTGTTTGGAGTATTGTGGTCCCTGCGTGTTCAGCTGCTGAGC 816
Db |||||
QY 697 AGTACTGGCCATACCTGTTTGGAGTATTGTGGTCCCTGCGTGTTCAGCTGCTGAGC 756
Db |||||
QY 817 CTTCCCTTCTCCCGAGAGCCACGCTACCTGCTCTTGGAGAGCACAAAGAGGCAAGA 876
Db |||||
QY 757 CTTCCCTTCTCCCGAGAGCCACGCTACCTGCTCTTGGAGAGCACAAAGAGGCAAGA 816
Db |||||
QY 877 GCTGTGAAAGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTTCCAAAGAGGTAGAGGAG 936
Db |||||
QY 817 GCTGTGAAAGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTTCCAAAGAGGTAGAGGAG 876
Db |||||
QY 937 GTCTGGCTGAGAGCCAGCTGAGAGGAGCATCCGCTGGTGTCCGTGCTGGAGCTGCTG 996
Db |||||
QY 877 GTCTGGCTGAGAGCCAGCTGAGAGGAGCATCCGCTGGTGTCCGTGCTGGAGCTGCTG 936
Db |||||
QY 997 AGAGCTCCCTAGCTCCGCTGGGAGTGTACCGTGTATTCACCATGGCCTGTACCAG 1056
Db |||||
QY 937 AGAGCTCCCTAGCTCCGCTGGGAGTGTACCGTGTATTCACCATGGCCTGTACCAG 996
Db |||||
QY 1057 CTCTGTGGCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGGAAAAGCTGGGATC 1116
Db |||||
QY 997 CTCTGTGGCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGGAAAAGCTGGGATC 1056
Db |||||
QY 1117 CCTCGGCAAGATCCCATACCTACCTTGTAGTAAGGGGGCATCGAGACTTTGGCTGCC 1176
Db |||||
QY 1057 CCTCTGGCAAGATCCCATACCTACCTTGTAGTAAGGGGGCATCGAGACTTTGGCTGCC 1116
Db |||||
QY 1177 GTCTTCTCTGGTTTGGTCAATTGAGACCTGGGAGGAGACCCCTCTCATTTGGTGGCTTT 1236
Db |||||
QY 1117 GTCTTCTCTGGTTTGGTCAATTGAGACCTGGGAGGAGACCCCTCTCATTTGGTGGCTTT 1176
Db |||||
QY 1237 GGGCTCATGGCTCTCTTTTGGGACCTTCACCATCAGCTGACCTGCAGGACCAAGCC 1296
Db |||||
QY 1177 GGGCTCATGGCTCTCTTTTGGGACCTTCACCATCAGCTGACCTGCAGGACCAAGCC 1236
Db |||||
QY 1297 CCTTGGCTCCCTTACCTGAGTATCGTGGGCAATTCCTGGCCATCATGCGCTCTTTCTGCAGT 1356
Db |||||
QY 1237 CCTTGGCTCCCTTACCTGAGTATCGTGGGCAATTCCTGGCCATCATGCGCTCTTTCTGCAGT 1296
Db |||||
QY 1357 GGGCCAGGTGGCATCCC 1373
Db |||||
QY 1297 GGGCCAGGTGTTTCCC 1313
Db |||||

Search completed: April 8, 2004, 11:57:02
Job time : 949 secs

[illegible]

```

PT /tag= a
PT /product= "Human TRICH-6 protein"
PT 100. .210
PT sig_peptide

```


PR 17-JUL-2001; 2001US-0306085P.
PR 24-JUL-2001; 2001US-0307536P.
PR 27-JUL-2001; 2001US-0308228P.
PR 30-JUL-2001; 2001US-0308877P.
PR 01-AUG-2001; 2001US-0309255P.
PR 17-AUG-2001; 2001US-0313328P.
PR 12-SEP-2001; 2001US-0318711P.
PR 19-SEP-2001; 2001US-0323380P.
PR 21-SEP-2001; 2001US-0323969P.
PR 04-JAN-2002; 2002US-0345022P.
PR 04-JAN-2002; 2002US-0345038P.
PR 28-FEB-2002; 2002US-0361172P.
PR 01-MAR-2002; 2002US-0360814P.
PR 01-MAR-2002; 2002US-0360830P.
PR 01-MAR-2002; 2002US-0361133P.
PR 01-MAR-2002; 2002US-0361147P.
PR 05-MAR-2002; 2002US-0361677P.
PR 02-APR-2002; 2002US-0363637P.
PR 12-APR-2002; 2002US-0372326P.
PR 16-APR-2002; 2002US-0372990P.
PR 19-APR-2002; 2002US-0373881P.
PR 19-APR-2002; 2002US-0373921P.
PR 02-JUL-2002; 2002US-00188186.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ;
PI Catterton E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L;
PI Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE;
PI Padigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S;
PI Shimkets RA, Spaderna SK, Spytsek KA, Stone DJ, Taupier RJ;
PI Vernet CAM, Voss EZ, Zhong M;
XX
XX WPI; 2003-221607/21.
DR P-FSDB; ADC26255.
XX
PT New isolated NOVX polypeptide, useful for determining the presence of, or
PT predisposition to a disease associated with altered levels of expression
PT of the polypeptide, and for treating or preventing cancer.
XX
PS Claim 20; SEQ ID NO 79; 478pp; English.
XX
XX The invention relates to a novel isolated NOV polypeptide. The
CC polypeptide of the invention demonstrates cytostatic activity and may be
CC used for determining the presence of, or predisposition to a disease
CC associated with altered levels of expression of the polypeptide,
CC including metabolic disorders, immune disorders, neurodegenerative
CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases
CC and cancer. The polypeptide may also be utilised during gene therapy
CC procedures, vaccine development and transgenic animal production. The
CC current sequence is that of the human NOV DNA of the invention.
XX
SQ Sequence 1502 BP; 329 A; 409 C; 401 G; 363 T; 0 U; 0 Other;
Query Match 52.5%; Score 1229; DB 9; Length 1502;
Best Local Similarity 88.2%; Pred. No. 4.6e-275;
Matches 1425; Conservative 0; Mismatches 10; Indels 180; Gaps 1;
Yy 60 GTCACTGAGCCCATGGCAAGGAAACAAATAGGAATTCAGGAATCGGGCTAGTTCC 119
Db 1 GTCACTGAGCCCATGGCAAGGAAACAAATAGGAATTCAGGAATCGGGCTAGTTCC 60
Yy 120 CCTCACAGATGACACCAAGCCACGCCGGCTCCAGGCCAGGGAGGCACTGCTGGAGTG 179
Db 61 CCTCACAGATGACACCAAGCCACGCCGGCTCCAGGCCAGGGAGGCACTGCTGGAGTG 120
Yy 180 TGACCACTGAGGAGTGGGTGCCAGGTGGAAGAGAGAAAGGACTGGTCTGCTCGCT 239
Db 121 TGTCCACTGAGGAGTGGGTGCCAGGTGGAAGAGAGAAAGGACTGGTCTGCTCGCT 180
Yy 240 CCTCGTGGCTCCCTCGGGGCGCTTCGGCTCCTCCTCCTCCTACGGCTACACCTGTC 299
Db 181 CCTCGTGGCTCCCTCGGGGCGCTTCGGCTCCTCCTCCTCCTCCTACGGCTACACCTGTC 240

QY 300 GGTGGTGAATGCCCCACCCCGTATCATCAAGCCCTTTTACAAATGAGTCATGGGAAAGAAG 359
Db 241 GGTGGTGAATGCCCCACCCCGTATCATCAAGCCCTTTTACAAATGAGTCATGGGAAAGAAG 300
QY 360 GCATGGACGTCAATAGACCCAGACACTCTGACCTGTCTCTGGTCTGTGACTGTGTCCAT 419
Db 301 GCATGGACGTCAATAGACCCAGACACTCTGACCTGTCTCTGGTCTGTGACTGTGTCCAT 360
QY 420 ATTCCGCATCGGTGACGTTGTGGGACGTTAATTGTGAAGATGATGGAAAGGTTCTTGG 479
Db 361 ATTCCGCATCGGTGACGTTGTGGGACGTTAATTGTGAAGATGATGGAAAGGTTCTTGG 420
QY 480 GAGNAGCACACTTGTGTGGCCAAATATGGTGTTCGCAATTTCTGCTGCAATTCGATTCGATGGC 539
Db 421 GAGNAGCACACTTGTGTGGCCAAATATGGTGTTCGCAATTTCTGCTGCAATTCGATTCGATGGC 480
QY 540 CTGCTCGCTCCAGCAGGAGCCTTTTGAATGCTCATTTGTGGGACGCTTCATCATCGGGCAT 599
Db 481 CTGCTCGCTCCAGCAGGAGCCTTTGAGATGCTCATTTGTGGGACGCTTCATCATCGGGCAT 540
QY 600 AGATGGAGCGCTCGCCCTCAGTGTGCTCCCATGCTAGCTCAGTGAGATCTCACCAAGGA 659
Db 541 AGATGGAGCGCTCGCCCTCAGTGTGCTCCCATGCTAGCTCAGTGAGATCTCACCAAGGA 600
QY 660 GATCCGTGGCTCTGTGGGACAGTGACGCCATCTTTATCTGCAATTTGGCGGTTCACGTGG 719
Db 601 GATCCGTGGCTCTGTGGGACAGTGACGCCATCTTTATCTGCAATTTGGCGGTTCACGTGG 660
QY 720 GCAGCTTCTGGGCTCGCCGAGCTGTGGGAAAGAGAGTACCTGGCCATACCTGTTTGG 779
Db 661 GCAGCTTCTGGGCTCGCCGAGCTGTGGGAAAGAGAGTACCTGGCCATACCTGTTTGG 720
QY 780 AGTGATTTGTGCTGCTGCGCTGTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGGCC 839
Db 721 AGTGATTTGTGCTGCTGCGCTGTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGGCC 780
QY 840 ACGTTACCTGCTTCTGGGAGAGCAACAGAGCGAAGAGCTGTGAAAGCTTCCAAACGTT 899
Db 781 ACGTTACCTGCTTCTGGGAGAGCAACAGAGCGAAGAGCTGTGAAAGCTTCCAAACGTT 840
QY 900 CTGGGTAAGCAGACGTTTCCAGAGGTAGAGAGGTCCTGCTGAGAGCCAGCTGCA 959
Db 841 CTGGGTAAGCAGACGTTTCCAGAGGTAGAGAGGTCCTGCTGAGAGCCAGCTGCA 900
QY 960 GAGGAGCATCCGCTGTGTCGCTGAGAGCTGCTGAGAGCTCCCTACGTCGCGTGCA 1019
Db 901 GAGGAGCATCCGCTGTGTCGCTGAGAGCTGCTGAGAGCTCCCTACGTCGCGTGCA 960
QY 1020 GGTGCTCACCGTGATTGTCAACATGGCCTGCTACAGCTCTGTGGCTCAATGCAATTG 1079
Db 961 GGTGCTCACCGTGATTGTCAACATGGCCTGCTACAGCTCTGTGGCTCAATGCAATTG 1020
QY 1080 GTTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATAGT 1139
Db 1021 GTTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATAGT 1080
QY 1140 CACTTGTAGTACAGGGGCACTGAGACTTTGGCTGCGCTCTTCTCTGTTGGTTCATTA 1199
Db 1081 CACTTGTAGTACAGGGGCACTGAGACTTTGGCTGCGCTCTTCTCTGTTGGTTCATTA 1125
QY 1200 GCACCTGGGACGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGG 1259
Db 1126 ----- 1125
QY 1260 GACCCTCACCATACGCTGAGCCCTGCAGGACCAAGCCCTGGGTCCCTTACCTAGTAT 1319
Db 1126 ----- 1125
QY 1320 COTGGGCATTCTGGCCATCATGCGCTCTTTCTGAGTGGGCGCAGGTGCGATCCCGTTTCAT 1379
Db 1126 -----TGGCATCCCGTTTCAT 1140

[REDACTED]

DB 958 TCTATACCAACAGCATCTTTGGAAAAGCTGGATCCCTCGGCAAGATCCCATAGTCA 1017
QY 1142 CTTGAGTACAGGGGACATCAGAGCTTTGGCTGCGTCTTCTCTGTTGGTCAATGAGC 1201
DB 1018 CTTGAGTACAGGGGACATCAGAGCTTTGGCTGCGTCTTCTCTGTTGGTCAATGAGC 1077
QY 1202 ACTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGA 1261
DB 1078 ACTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGA 1137
QY 1262 CCTCACCATCAGCTGACCTCTGAGGACCAACGCGCCCTGGTCCCTACTGAGTATCG 1321
DB 1138 CCTCACCATCAGCTGACCTCTGAGGACCAACGCGCCCTGGTCCCTACTGAGTATCG 1197
QY 1322 TGGGCAATCTGGCCATCATCGCTCTTCTGAGTGGGCGCAGGTGGGATCCGCTTCACT 1381
DB 1198 TGGGCAATCTGGCCATCATCGCTCTTCTGAGTGGGCGCAGGTGGGATCCGCTTCACT 1257
QY 1382 TGACTGGTGGTCTTCCAGCAATCTCAGGGCGCGCTCTTCACTATTCAGGACCG 1441
DB 1258 TGACTGGTGGTCTTCCAGCAATCTCAGGGCGCGCTCTTCACTATTCAGGACCG 1317
QY 1442 TCAACTGGCTCTCCAACTTTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCTGG 1501
DB 1318 TCAACTGGCTCTCCAACTTTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCTGG 1377
QY 1502 ACACCTACTGTTCTAGTCTTCTGCTACAAATTTGTATCAGAGTGTCTACTGTAT 1561
DB 1378 ACACCTACTGTTCTAGTCTTCTGCTACAAATTTGTATCAGAGTGTCTACTGTAT 1437
QY 1562 TTGTGCTGCTGAGACCAAAACAGAACCTATGAGAAATCAGCAAAATCAGCAGGATTTTCCAAA 1621
DB 1438 TTGTGCTGCTGAGACCAAAACAGAACCTATGAGAAATCAGCAAAATCAGCAGGATTTTCCAAA 1497
QY 1622 GGAACAAGACATACCCACGAGAGAGAAATCGACTGCTGCTGATGCT 1674
DB 1498 GGAACAAGACATACCCACGAGAGAGAAATCGACTGCTGCTGATGCT 1550

RESULT 11

ADB48664
ID ADB48664 standard; cDNA; 2842 BP.

AC ADB48664;

DT 04-DEC-2003 (first entry)

DE Novel human cDNA SEQ ID NO 574.

KW ss; cancer; neurodegenerative disease; human.

OS Homo sapiens.

PN US2003104529-A1.

PD 05-JUN-2003.

PF 04-JAN-2002; 2002US-00037270.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 19-JUL-2000; 2000US-00620312.

PA (ZHOU/) ZHOU P.

PA (TANG/) TANG Y T.

PA (LIUC/) LIU C.

PA (ASUN/) ASUNDI V.

PA (DRMA/) DRMANAC R T.

XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;

XX WPI; 2003-678194/64.

XX PT New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX Claim 1; SEQ ID NO 574; 99pp; English.
XX The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.

XX SQ Sequence 2842 BP; 698 A; 709 C; 693 G; 733 T; 0 U; 9 Other;

Query Match 56.2%; Score 1317.8; DB 8; Length 2842;

Best Local Similarity 89.9%; Pred. No. 1.4e-295;

Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;

QY 2 CGACCCACGGCTTGGCAGAGTCTGGGCTCCCTGGAGCTGAGCCATCAGCTGGGT 61
DB 40 CGACCCACGGCTGCG-CCCTTGGCAGAGTCTGGGCTCCCTGGAGCTGAGCCATCAGCTGGGT 98
QY 62 CACTGAGACCCATGGCAGGAAACAAATAGGAATTCAGGAGTGGGCTAGTCCGC 121
DB 99 CACTGAGACCCATGGCAGGAAACAAATAGGAATTCAGGAGTGGGCTAGTCCGC 158
QY 122 TCACAGATGACACAGCCACGCGGCTCCAGGGCCAGGAGGACACTGCTGAGTGTG 181
DB 159 TCACAGATGACACAGCCACGCGGCTCCAGGGCCAGGAGGACACTGCTGAGTGTG 218
QY 182 ACCACTGAGAGTGGGCTGCCAGGTGGAAGGAGAAAGGACTGGTCTGCTGCTCC 241
DB 219 ACCACTGAGAGTGGGCTGCCAGGTGGAAGGAGAAAGGACTGGTCTGCTGCTCC 278
QY 242 TCCTGGGCTCCCTGGGCGGCTTGGGCTCCCTCTCTACGCTGACACTGCTGCG 301
DB 279 TCGTGGGCTCCCTGGGCGGCTTGGGCTCCCTCTCTACGCTGACACTGCTGCG 338
QY 302 TGGTGAATGCCCCACCCGCTACATCAAGGCGCTTTTACAATGAGTCAATGGAAAGAGGC 361
DB 339 TGGTGAATGCCCCACCCGCTTTTACAATGAGTCAATGGAAAGAGGC 357
QY 362 ATGAGCTGCCAATAGACCCAGACACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCT 421
DB 358 ----- 357
QY 422 TCGCCATCGGTGACTTGTGGGACGTTAATTTGAGATGATTTGAAAGGTTCTTGGGA 481
DB 358 ----- 357
QY 482 GGAAGCACACTTTGCTGGGCAATAATGGGTTTGGCAATTTCTGCTGCTGCTGCTGCTGCT 541
DB 358 GGAAGCACACTTTGCTGGGCAATAATGGGTTTGGCAATTTCTGCTGCTGCTGCTGCTGCT 417
QY 542 GCTGCTTCCAGGAGGAGCTTTTGAATACTCATTTGAGGAGCGCTTTCATCATGGGCGATAG 601
DB 418 GCTGCTTCCAGGAGGAGCTTTTGAATACTCATTTGAGGAGCGCTTTCATCATGGGCGATAG 477
QY 602 ATGAGGCGTTCGCGCTCAGTGTGCTCCCATGATGATGATGATGATGATGATGATGATGAT 661
DB 478 ATGAGGCGTTCGCGCTCAGTGTGCTCCCATGATGATGATGATGATGATGATGATGATGAT 537
QY 662 TCGTGGCTCTTGGGCGAGGTGACTGCGCATCTTTATCTGCAATGGGCTGTTCACTGGGC 721
DB 538 TCGTGGCTCTTGGGCGAGGTGACTGCGCATCTTTATCTGCAATGGGCTGTTCACTGGGC 597
QY 722 AGCTTCTGGGCTGCGCGAGCTGCTGGGAAAGAGAGTACCTGGGCAATACCTGTTTGGAG 781
DB 598 AGCTTCTGGGCTGCGCGAGCTGCTGGGAAAGAGAGTACCTGGGCAATACCTGTTTGGAG 657

2	CGACCCACGGCTCGGGCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT	61	QY
40	CGACCCACGGGTGC - GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT	98	DB
62	CACCTGAGACCCATCGGCAAGAAACAAAATAGGAATTCCAAGAACTGGGCCCTAGTTCGCC	121	QY
99	CACCTGAGACCCATCGCAGAGAAACAAAATAGGAATTCCAAGGAATCGGCCCTAGTTCGCC	158	DB
122	TCACAGATGACACACAGCCACGCCCGGCCCTCCAGGGCCAGGGAGGGGCACTGCTGGAGTGTG	181	QY
159	TCA CAGATGACACACAGCCACGCCCAGGCCCTCCAGGGCCAGGGAGGGGCACTGCTGGAGTGTG	218	DB
182	ACCACCTGAGAGTGGGGTSCCAGGTGGAAGAGAGAAAGAGACTGGTCTGCTCGCTCC	241	QY
219	ACCACCTGAGAGTGGGGTSCCAGGTGGAAGAGAGAAAGAGACTGGTCTGCTCGCTCC	278	DB
242	TCGTGGCCTCCCTCGGGGGCGCTTCGGCTCCCTCTCTCTACGGCTACAACCTGTGCG	301	QY
279	TCGTGGCCTCCCTCGGGGGCGCTTCGGCTCCCTCTCTCTACGGCTACAACCTGTGCG	338	DB
302	TGGTGAATGCCCCCACCCTGATCAAGGCCCTTTTACAATGAGTCATGGGAAGAAGAGC	361	QY
339	TGGTGAATGCCCCCACCCTGATCAAGGCCCTTCCTGACTCTGCTCTGTGACTGTGCCATAT	421	QY
358	-----	357	DB
422	TCGCCATCGGTGACCTTGTGGGAGCGTTAAATTGTGAAGATGATTGGAAGGTTCTTGGGA	481	QY
358	-----	357	DB
482	GGAGACACCTTTGCTGGCCATAATAGGGTTTGCAATTTCTGCTGATTCGTATGGCCT	541	QY
358	GGAGACACCTTTGCTGGCCATAATAGGGTTTGCAATTTCTGCTGATTCGTATGGCCT	417	DB
542	GCTCGCTCAGGCAGAGAGCCTTTGAAATGCTCATTTGTGGGAGCGTTTCATCGGCCATAG	601	QY
418	GCTCGCTCAGGCAGAGAGCCTTTGAAATGCTCATCGTGGGAGCGTTTCATCGGCCATAG	477	DB
602	ATGAGAGCGCTCGCCCTCAGTGTGCTCCCACTGATCCTCAGTCAGATCTCACCCCAAGGAGA	661	QY
478	ATGAGAGCGCTCGCCCTCAGTGTGCTCCCACTGATCCTCAGTCAGATCTCACCCCAAGGAGA	537	DB
662	TCGTGGCCTCTGGGGCAGGTGACTGCCATCTTTATCTGCATTCGGGTTCATCTGGGC	721	QY
538	TCGTGGCCTCTGGGGCAGGTGACTGCCATCTTTATCTGCATTCGGGTTCATCTGGGC	597	DB
722	AGCTTCTGGCCTGCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAG	781	QY
598	AGCTTCTGGCCTGCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAG	657	DB
782	TGATTGTGTCCTTGGCCTGTCCTGAGCTCTGAGCCTTCCCTTTCTCCGGACAGCCCAAC	841	QY
658	TGATTGTGTCCTTGGCCTGTCCTGAGCTGTCAGCCTTCCCTTTCTCCGGACAGCCCAAC	717	DB
842	GCTACCTGCTCTTGGAGAGACACACGAGGGCAGAGCTGTGAAAGCCTTCCAAAGTTCCT	901	QY
718	GCTACCTGCTCTTGGAGAGACACACGAGGGCAGAGCTGTGAAAGCCTTCCAAAGTTCCT	777	DB
902	TGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGAGGTCTGTGCTCAGAGCCACGCTGACAG	961	QY
778	TGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGAGGTCTGTGCTCAGAGCCACGCTGACAG	837	DB
962	GGAGCATCCGCCCTGGTGTCCGTCGTGGAGCTGCTGAGAGCTCCCTTACGTCGGCTGGCAGG	1021	QY
838	GGAGCATCCGCCCTGGTGTCCGTCGTGGAGCTGCTGAGAGCTCCCTTACGTCGGCTGGCAGG	897	DB
1022	TGGTCAACCGTATTGTACCATGGCTGCTACAGCTCTGTGGCCTCAATGCAATTTGGT	1081	QY
898	TGGTCAACCGTATTGTACCATGGCTGCTACAGCTCTGTGGCCTCAATGCAATTTGGT	957	DB
1082	TCTATACCAACAGCATCTTTGGAAAAGCTGGGATCCCTCCGGCAAGATCCCATACGTCA	1141	QY

RESULT 10
 AAI58686
 AAI58686 standard; cDNA; 2842 BP.
 XX
 AAI58686;
 XX
 22-OCT-2001 (first entry)
 XX
 Human polynucleotide SEQ ID NO 889.
 XX
 Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia; ss.
 XX
 Homo sapiens.
 XX
 WO200153312-A1.
 XX
 26-JUL-2001.
 XX
 26-DEC-2000; 2000WO-US034263.
 XX
 23-DEC-1999; 99US-00471275.
 XX
 21-JAN-2000; 2000US-00488725.
 XX
 25-APR-2000; 2000US-00552317.
 XX
 20-JUN-2000; 2000US-00598042.
 XX
 19-JUL-2000; 2000US-00620312.
 XX
 03-AUG-2000; 2000US-00653450.
 XX
 14-SEP-2000; 2000US-00662191.
 XX
 19-OCT-2000; 2000US-00693036.
 XX
 29-NOV-2000; 2000US-00727344.
 XX
 (HYSE-) HYSEQ INC.
 XX
 Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 Zhou P, Goodrich R, Drmanac RT;
 WPI; 2001-442253/47.
 XX
 P-FSDB; AAM39530.
 XX
 Novel nucleic acids and polypeptides, useful for treating disorders such
 as central nervous system injuries.
 XX
 Claim 1; SEQ ID NO 889; 10078pp; English.
 XX
 The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 immunosuppressant and cytostatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 localised neuropathies and central nervous system diseases, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 utilisation of the activities such as: Immune system suppression,
 Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 assays for receptor activity, arthritis and inflammation, leukaemias and
 C.N.S disorders. Note: The sequence data for this patent did not form
 part of the printed specification
 XX
 Sequence 2842 BP; 698 A; 709 C; 693 G; 733 T; 0 U; 9 Other;
 Query March 56.2%; Score 1317.8; DB 4; Length 2842;
 Best Local Similarity 89.9%; Pred. No. 1.4e-295;
 Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps

5Q	Sequence	1710 BP; 393 A; 444 C; 442 G; 431 T; 0 U; 0 Other;
	Query Match	59.7%; Score 1399; DB 9; Length 1710;
	Best Local Similarity	93.6%; Pred. No. 1.8e-314;
	Matches 1511; Conservative	0; Mismatches 5; Indels 99; Gaps 1;
QY	60	GTCACTGAGACCCATGCGAAGAAACAAATAGGAATCCAAAGAACTGGGCTAGTTCC 1119
DB	1	GTCACTGAGACCCATGCGAAGAAACAAATAGGAATCCAAAGAACTGGGCTAGTTCC 60
QY	120	CCTCAGACATGACACCCAGCCAGCGCGGCCCTCCAGGCGCAGGAGGAGCACTGCTGGAGTG 179
DB	61	CCTCAGACATGACACCCAGCCAGCGCGGCCCTCCAGGCGCAGGAGGAGCACTGCTGGAGTG 120
QY	180	TGACCACCTGAGGAGTGGGGTCCAGGTGGAGGAGGAGAAAGCACTGGTCTGCTCGCT 239
DB	121	TGACCACCTGAGGAGTGGGGTCCAGGTGGAGGAGGAGAA----- 162
QY	240	CCTGCTGCGCTCCCTCGCGGGGCGCTTCGGGCTCCTCTCTCTACGGCTACAACTGTC 299
DB	163	----- 162
QY	300	GGTGTGAATGCCCCCACCCTGATCAAGGCGCTTTTACAAATGAGTCATGGGAAGAAG 359
DB	163	-----GTACATCAAGGCGCTTTTACAAATGAGTCATGGGAAGAAG 201
QY	360	GCATGGACGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTGCTCAT 419
DB	202	GCATGGACGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTGCTCAT 261
QY	420	ATTGCCCATCGTGGACTTTGGGGAAGTTAAATTTGAAGATGATGGAAAGTTCTTGG 479
DB	262	ATTGCCCATCGTGGACTTTGGGGAAGTTAAATTTGAAGATGATGGAAAGTTCTTGG 321
QY	480	GAGGAAGACACTTTGCTGGGCAATATATGGTTTGGCAATTTCTGCTGCTGCTGCTGCTGCTG 539
DB	322	GAGGAAGACACTTTGCTGGGCAATATATGGTTTGGCAATTTCTGCTGCTGCTGCTGCTGCTG 381
QY	540	CTGCTCGCTCCAGGCGAGAGCTTTGAAATGCTCATTTGCTGGAGCGCTTCATCATGAGGCAT 599
DB	382	CTGCTCGCTCCAGGCGAGAGCTTTGAAATGCTCATTTGCTGGAGCGCTTCATCATGAGGCAT 441
QY	600	AGATGGAGGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGA 659
DB	442	AGATGGAGGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGA 501
QY	660	GATCGTGGCTCTCTGGGGGAGGTGACTGCAATCTTTATCTGCATTTGGCGTGTCTACTGG 719
DB	502	GATCGTGGCTCTCTGGGGGAGGTGACTGCAATCTTTATCTGCATTTGGCGTGTCTACTGG 561
QY	720	GCAGCTTCTGGGCTGCCGAGCTGCTGGGAAGGAGAGTACCTGGCCATACCTGTTGG 779
DB	562	GCAGCTTCTGGGCTGCCGAGCTGCTGGGAAGGAGAGTACCTGGCCATACCTGTTGG 621
QY	780	AGTGATTTGGTCCCTGCGCTGTCCAGTGTGCTGAGCTTCCCTTTCTCCCGGACAGGCC 839
DB	622	AGTGATTTGGTCCCTGCGCTGTCCAGTGTGCTGAGCTTCCCTTTCTCCCGGACAGGCC 681
QY	840	ACGCTACCTGCTCTTTGGGAAGCACACGAGGACAGAGCTGTGAAGCCTTTCCAAACGTT 899
DB	682	ACGCTACCTGCTCTTTGGGAAGCACACGAGGACAGAGCTGTGAAGCCTTTCCAAACGTT 741
QY	900	CTTGGGTAAAGCAGACGTTTCCCAAGAGTAGAGAGGTCTCTGGCTGAGAGCGCGTGCA 959
DB	742	CTTGGGTAAAGCAGACGTTTCCCAAGAGTAGAGAGGTCTCTGGCTGAGAGCGCGTGCA 801
QY	960	GAGGAGCATCCGCTGGTGTCCGCTGAGAGTGTGAGAGTCCCTACCTGCTCCCTGGCA 1019
DB	802	GAGGAGCATCCGCTGGTGTCCGCTGAGAGTGTGAGAGTCCCTACCTGCTCCCTGGCA 861
QY	1020	GTGGGTACCGGTATTTGTCACCATGCGCTTACAGCTCTGTGGCTTCAATGCAATTTG 1079
DB	862	GTGGGTACCGGTATTTGTCACCATGCGCTTACAGCTCTGTGGCTTCAATGCAATTTG 921

QY	1080	GTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCGGCAAAAGATCCCATACGT 1139
DB	922	GTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCTGGCAAAAGATCCCATACGT 981
QY	1140	CACCTTGAAGTACAGGGGGCATCGAGACTTTGGCTGCCGTCTTCTCTGTTGGTTCATTGA 1199
DB	982	CACCTTGAAGTACAGGGGGCATCGAGACTTTGGCTGCCGTCTTCTCTGTTGGTTCATTGA 1041
QY	1200	GACCTGGGACGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGG 1259
DB	1042	GACCTGGGACGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGG 1101
QY	1260	GACCTCTACCATCAACGCTGACCTCGAGACCAACGCCCTGGGTCCTTCTTCTTCTTCT 1319
DB	1102	GACCTCTACCATCAACGCTGACCTCGAGACCAACGCCCTGGGTCCTTCTTCTTCTTCT 1161
QY	1320	CGTGGGCAATCTGGGCAATCATCGCTCTTCTGAGTGGGCGAGGTGGCATCCCGTTTCT 1379
DB	1162	CGTGGGCAATCTGGGCAATCATCGCTCTTCTGAGTGGGCGAGGTGGCATCCCGTTTCT 1221
QY	1380	CTTGAATGCTGAGTCTTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTTCTTCTTCTTCT 1439
DB	1222	CTTGAATGCTGAGTCTTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTTCTTCTTCTTCT 1281
QY	1440	CGTCAACTGGCTCTCCAACTTTGCTGGTGGTCTCTTCCCATTCATTCAGAAAGTCT 1499
DB	1282	CGTCAACTGGCTCTCCAACTTTGCTGGTGGTCTCTTCCCATTCATTCAGAAAGTCT 1341
QY	1500	GGACACCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1559
DB	1342	GGACACCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1401
QY	1560	TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCAGGCAATTTTCCAA 1619
DB	1402	TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCAGGCAATTTTCCAA 1461
QY	1620	AAGGAAACAAAGCATATCCCAAGAGAGAGAAATCGACTCAGTGTGCTGATGCT 1674
DB	1462	AAGGAAACAAAGCATATCCCAAGAGAGAGAAATCGACTCAGTGTGCTGATGCT 1516
RESULT 9		
AAKS1571		
ID	AAKS1571	standard; cDNA; 1704 BP.
XX	AAKS1571;	
AC	AAKS1571;	
XX	06-NOV-2001	(first entry)
DT	06-NOV-2001	(first entry)
XX	Human polynucleotide	SEQ ID NO 116.
DE	Human; cytokines; cell proliferation; cell differentiation; gene therapy;	
XX	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation; ss.	
XX	Homo sapiens.	
OS	WO200157190-A2.	
PN	09-AUG-2001.	
XX	05-FEB-2001; 2001WO-US004098.	
PF	03-FEB-2000; 2000US-00496914.	
XX	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	

QY	204	AGTGGAGAGAGAGAAAGAGCTGGTCTCTGCTCTCTCGTGGCTCTCCCTCGGGGCGC	263
Db	188	AGGTGGAAGAGAGAGAAAGGACCTGGTCTCTGCTCTCTCGTGGCTCTCCCTCGGGGCGC	247
QY	264	CTTTCGGCTCTCTCTCTCTCTACGGCTACACCTGTGGTGGTGAATGCCCCACCCCGTA	323
Db	248	CTTTCGGCTCTCTCTCTCTACGGCTACACCTGTGGTGGTGAATGCCCCACCCCGTA	307
QY	324	CATCAAGGCGCTTTTCAATGAGTCATGGGAAAGGACATGGACGTCCAAATAGACCCAGA	383
Db	308	CATCAAGGCGCTTTTCAATGAGTCATGGGAAAGGACATGGACGTCCAAATAGACCCAGA	367
QY	384	CACCTCTGACTCTGCTCTCTGGTCTGTGATGTGTCCATATTCGCCCATCGGTGGACTTTGGG	443
Db	368	CACCTCTGACTCTGCTCTCTGGTCTGTGATGTGTCCATATTCGCCCATCGGTGGACTTTGGG	427
QY	444	GACGTTAAATTGTGAAGATGATTTGGAAGGTTCTTGGGAGGAAGACACACTTTGCTGGCCAA	503
Db	428	GACATTTAAATTGTGAAGATGATTTGGAAGGTTCTTGGGAGGAAGACACACTTTGCTGGCCAA	487
QY	504	TAAATGGGTTTGCAATTTCTGCTGCATGTGATGGCCTGCTCGCTCAGCAGAGAGCGTT	563
Db	488	TAAATGGGTTTGCAATTTCTGCTGCATGTGATGGCCTGCTCGCTCAGCAGAGAGCGTT	547
QY	564	TGAAATGCTCATTTGTGGGAGCGTTTCATCATGCGCATAGATGGAGGCGTCGCCCTCAGTGT	623
Db	548	TGAAATGCTCATCGTGGGAGCGTTTCATCATGCGCATAGATGGAGGCGTCGCCCTCAGTGT	607
QY	624	GCTCCCCATGTACCTCAGTGAGATCTCACCCAAAGGAGATCCGTGGCTCTCTGGGGCAGGT	683
Db	608	GCTCCCCATGTACCTCAGTGAGATCTCACCCAAAGGAGATCCGTGGCTCTCTGGGGCAGGT	667
QY	684	GACTGGCCATCTTTATCTGCATTTGGCGTGTTCACTGGGCAGCTCTGGGCGTGCOCGAGCT	743
Db	668	GACTGGCCATCTTTATCTGCATTTGGCGTGTTCACTGGGCAGCTCTGGGCGTGCOCGAGCT	727
QY	744	GCTGGGAAAGGAGTACCTGGCCATACCTGTTTGGAGTGATTTGTGTCCTCGCGTTGT	803
Db	728	GCTGGGAAAGGAGTACCTGGCCATACCTGTTTGGAGTGATTTGTGTCCTCGCGTTGT	787
QY	804	CGAGCTGTGAGCGCTTCCCTTTCTCCGGAGACGCCACGCTACCTGCTCTTGGAGAGCA	863
Db	788	CGAGCTGTGAGCGCTTCCCTTTCTCCGGAGACGCCACGCTACCTGCTCTTGGAGAGCA	847
QY	864	CAACGAGGCAAGAGCTGTGAAAGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTCCCA	923
Db	848	CAACGAGGCAAGAGCTGTGAAAGCCTTCCAAAGCTTCTGGGTAAAGCAGACGTTCCCA	907
QY	924	AGAGGTAGAGGAGGTCTGTGCTGAGAGCCACGTGCAGAGGAGCATCCGCTGTGTCGGT	983
Db	908	AGAGGTAGAGGAGGTCTGTGCTGAGAGCCGCTGAGAGGAGCATCCGCTGTGTCGGT	967
QY	984	GCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAGGTGGTCAACGTGATGTACCAT	1043
Db	968	GCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAGGTGGTCAACGTGATGTACCAT	1027
QY	1044	GGCGCTGCTACAGCTCTGTGGCGCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGG	1103
Db	1028	GGCGCTGCTACAGCTCTGTGGCGCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGG	1087
QY	1104	AAAAGCTGGATCCCTCCGGCAAGATCCCATACGTCACTTGTAGTACAGGGGGCATCGA	1163
Db	1088	AAAAGCTGGATCCCTCCGGCAAGATCCCATACGTCACTTGTAGTACAGGGGGCATCGA	1147
QY	1164	GACTTTGGCTGCGCTCTCTCTGTTTGGTCAATTTGAGCACCTGGGACCGGAGACCCCTCCT	1223
Db	1148	GACTTTGGCTGCGCTCTCTCTGTTTGGTCAATTTGAGCACCTGGGACCGGAGACCCCTCCT	1207
QY	1224	CATTGTGGCTTTGGGCTCATGGGCTCTTCTTTTGGGACCTTCAACCATCAAGTGGACCT	1283
Db	1208	CATTGTGGCTTTGGGCTCATGGGCTCTTCTTTTGGGACCTTCAACCATCAAGTGGACCT	1267
QY	1284	GCAGGACACGCCCCCTTGGTCCCTTACCTGAGTATCGTGGGCAATCTTGGGCAATCATGCG	1343

QY 497 TGGCCAAATATGGGTTTGCATTTCTGCTGATTTGCTGATGGCTGCTGCTGCCAGGAG 556
Db 481 TGGCCAAATATGGGTTTGCATTTCTGCTGATTTGCTGATGGCTGCTGCTGCCAGGAG 540
QY 557 GAGCCTTTGAAATGCTCATTTGTTGGAGCGCTTTCATCATGGCCATAGATGAGGCGTGCCTCC 616
Db 541 GAGCCTTTGAAATGCTCATCGTGGAGCGCTTTCATCATGGCCATAGATGAGGCGTGCCTCC 600
QY 617 TCAGTGTGCTCCCATGCTACCTAGTGAATCTCACCAGGAGATCCGTGGCTCTCTGG 676
Db 601 TCAGTGTGCTCCCATGCTACCTAGTGAATCTCACCAGGAGATCCGTGGCTCTCTGG 660
QY 677 GGCAGGTGACCTGCTCATTTATCTGCAATTTGGCGTGTTCACCTGGCGAGCTTCTGGGCTGC 736
Db 661 GGCAGGTGACCTGCTCATTTATCTGCAATTTGGCGTGTTCACCTGGCGAGCTTCTGGGCTGC 720
QY 737 CCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATTTGGTCCCTG 796
Db 721 CCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATTTGGTCCCTG 780
QY 797 CCGTGTCCAGCTGCTGAGCTTCCCTTTCTCCGGACAGCCACAGCTACCTGCTCTTGG 856
Db 781 CCGTGTCCAGCTGCTGAGCTTCCCTTTCTCCGGACAGCCACAGCTACCTGCTCTTGG 840
QY 857 AGAAGCACAAGGAGGAGAGCTGCTGAAAGCTTCCAAAGCTTCTGGGTAAAGCAGAGC 916
Db 841 AGAAGCACAAGGAGGAGAGCTGCTGAAAGCTTCCAAAGCTTCTGGGTAAAGCAGAGC 900
QY 917 TTTCCCAAGAGGTAGAGAGTCTGCTGAGAGCCAGCTGAGAGAGCATCCGCTGCTG 976
Db 901 TTTCCCAAGAGGTAGAGAGTCTGCTGAGAGCCAGCTGAGAGAGCATCCGCTGCTG 960
QY 977 TGTCCGTGCTGAGCTGCTGAGAGCTCCCTAGCTCCGCTGGCAGGTGGTCACTGATTTG 1036
Db 961 TGTCCGTGCTGAGCTGCTGAGAGCTCCCTAGCTCCGCTGGCAGGTGGTCACTGATTTG 1020
QY 1037 TCACCATGCTGCTTACAGCTGCTGGCTCAATGCAATTTGTTCTATACCAAGCA 1096
Db 1021 TCACCATGCTGCTTACAGCTGCTGGCTCAATGCAATTTGTTCTATACCAAGCA 1080
QY 1097 TCTTTGAAAAAGCTGGGATCCCTCCGCAAGATCCCATACGTCACCTTGAGTACAGGGG 1156
Db 1081 TCTTTGAAAAAGCTGGGATCCCTCCGCAAGATCCCATACGTCACCTTGAGTACAGGGG 1140
QY 1157 GCATCGAGACTTTGGCTGCGCTTCTCTGGTTTGGTCAATGAGACCTGGGAGGAGAC 1216
Db 1141 GCATCGAGACTTTGGCTGCGCTTCTCTGGTTTGGTCAATGAGACCTGGGAGGAGAC 1200
QY 1217 CCCTCCTCATTTGGTGGCTTGGGCTCATGGGCTTCTTTTGGGACCTCACCATCACGC 1276
Db 1201 CCCTCCTCATTTGGTGGCTTGGGCTCATGGGCTTCTTTTGGGACCTCACCATCACGC 1260
QY 1277 TGACCTGSCAGGACCAAGCCCTGGGTCCCTTACCTGAGTATCGTGGGATTTCTGGCA 1336
Db 1261 TGACCTGSCAGGACCAAGCCCTGGGTCCCTTACCTGAGTATCGTGGGATTTCTGGCA 1320
QY 1337 TCATGCGCTCTTTCTGAGTGGGCGAGTGGCATCCCGTTTCATTTGACTGGTGGTCTT 1396
Db 1321 TCATGCGCTCTTTCTGAGTGGGCGAGTGGCATCCCGTTTCATTTGACTGGTGGTCTT 1380
QY 1397 TCCAGCAATCTCAGCGGCGGCTGCCCTTTCATCATTTGAGGACCGTCAACTGGCTCTCA 1456
Db 1381 TCCAGCAATCTCAGCGGCGGCTGCCCTTTCATCATTTGAGGACCGTCAACTGGCTCTCA 1440
QY 1457 ACTTTGCTGTTGGGCTCTCTTCCCATTCATTCAGAAAAGCTTGGACACTACTGTTTCC 1516
Db 1441 ACTTTGCTGTTGGGCTCTCTTCCCATTCATTCAGAAAAGCTTGGACACTACTGTTTCC 1500
QY 1517 TAGTCTTTGCTACAAATTTGATATCAGAGTGTATCTACCTGATTTTGTGCTGCTGAGA 1576
Db 1501 TAGTCTTTGCTACAAATTTGATATCAGAGTGTATCTACCTGATTTTGTGCTGCTGAGA 1560

QY 1577 CCAAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCCAAAAGAAACAAAGCATACC 1636
Db 1561 CCAAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCCAAAAGAAACAAAGCATACC 1620
QY 1637 CACCAGAGAGAAATCGACTCAGCTGCTGCTGATGCTC 1675
Db 1621 CACCAGAGAGAAATCGACTCAGCTGCTGCTGATGCTC 1659

RESULT 5

ABA09154

ID ABA09154 standard; cDNA; 1851 BP.

XX ABA09154;

AC ABA09154;

DT 11-JAN-2002 (first entry)

XX Human GLUT9 homologue-encoding cDNA, SEQ ID NO: 930.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antiaesthetic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnerary; antiulcer; ss.

OS Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US003800.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX P-PSDB; ABB11910.

XX Human proteins and DNA encoding sequences useful for preventing, treating
or ameliorating a medical condition in a mammalian subject e.g. arthritis
and cancer.

XX Claim 1; Page 802; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin or inhibin-related activities;
XX chemoattractant or chemokinetic activities; haemostatic; thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be

XX	(INCY-) INCYTE GENOMICS INC.
PA	
XX	
PI	Thornton M, Wallia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;
PI	Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YT;
PI	Azinzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BB;
PI	Lee EA, Hafalia A, Baughn WR, Green BD, Khan FA, Kearney L;
PI	Elliot VS, Seilmaner JJ, Policky JL, Borowsky ML, Burford N, Ding L;
PI	Lu DAM, Hillman JL;
XX	
DR	WPI; 2002-122055/16.
DR	P-PGDB; AME16788.
XX	
PT	New human transporters and ion channels (TRICH) polypeptides useful for
PT	diagnosing, treating or preventing disorders associated with aberrant
PT	expression of TRICH.
XX	
PS	Claim 11; Page 207-208; 210pp; English.
XX	
CC	The invention relates to human transporters and ion channels (TRICH)
CC	polypeptides and their cDNA molecules. The nucleic acid and polypeptide
CC	sequences are useful in the diagnosis, treatment, and prevention of
CC	disorders associated with transport (akinesia, cystic fibrosis, Bell's
CC	palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,
CC	amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's
CC	muscular dystrophy); immunological (AIDS, Addison's disease, allergies,
CC	asthma); cell proliferative disorders (cancers, leukamia, porriasis);
CC	cardiac disease (angina, hypertension, or bradyarrhythmia) and in the
CC	assessment of the effects of exogenous compounds on the expression of
CC	nucleic acid and amino acid sequences of transporters and ion channels.
CC	The polynucleotides may be used to detect and quantify gene expression in
CC	biopsied tissues in which TRICH expression may be correlated with a
CC	disease, to generate hybridization probes for mapping naturally occurring
CC	genomic sequence, and in drug screening. The present sequence is human
CC	TRICH-25 cDNA
XX	
SQ	Sequence 2032 BP; 438 A; 557 C; 536 G; 501 T; 0 U; 0 Other;
	Query Match 70.5%; Score 1651; DB 6; Length 2032;
	Best Local Similarity 99.7%; Pred. No. 0;
	Matches 1654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	17 GCCTTGGCAGAGTCGTGGGTCCCTGCAGCTGAGGCCATCAGCTGGTGCTACGTAGACCATTGG 76
DB	1 GCCTTGGCAGAGTCGTGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCATTGG 60
QY	77 CAAGGAACAATAATAGGAATTCGAAGAACTGGGCTAGTTCCCTCTCACAGATGACACCA 136
DB	61 CAAGGAACAATAATAGGAATTCGAAGAACTGGGCTAGTTCCCTCTCACAGATGACACCA 120
QY	137 GCCACGCGGGGCTCCAGGCCCAGGAGGACCATCTCTGGAGTGTACCACTGAGAGTGG 196
DB	121 GCCACGCGGGGCTCCAGGCCCAGGAGGACCATCTCTGGAGTGTACCACTGAGAGTGG 180
QY	197 GGGTGCCAGGTGGAGGAGAAGAACGATGGTCTCTCTCGCTCTCTCGCTGCTCCCTCC 256
DB	181 GGGTGCCAGGTGGAGGAGAAGAACGATGGTCTCTCTCGCTCTCTCGCTGCTCCCTCC 240
QY	257 CGGCGCGCTTCGGGCTCCCTCCCTCTACGGGTCAAACTGTCGGTGGTGAATGCCGCCCA 316
DB	241 CGGCGCGCTTCGGGCTCCCTCCCTCTACCGGCTACAACTGTCGGTGGTGAATGCCGCCCA 300
QY	317 CCCGTHCATCAAGGCCCTTTTCAATGATGTCATGGGAAGAAGGCGATCGAGCTCCAATAG 376
DB	301 CCCGTHCATCAAGGCCCTTTTCAATGATGTCATGGGAAGAAGGCGATCGAGCTCCAATAG 360
QY	377 ACCCAGACACTCTGACCTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB	361 ACCCAGACACTCTGACCTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY	437 TTGTGGGGACGTTAAATTTGTGAAGATGATTTGGAAAAGTTCTTTGGGAGGAGACACACTTTGC 496
DB	421 TTGTGGGGACGTTAAATTTGTGAAGATGATTTGGAAAAGTTCTTTGGGAGGAGACACACTTTGC 480

841	DB	C	G	T	A	C	T	G	C	T	T	T	G	G	A	G	A	G	C	A	C	A	A	G	G	G	C	A	G	A	G	C	T	G	T	G	A	A	G	C	T	T	C	C	A	A	G	T	C		900	
901	QY	T	T	G	G	T	A	A	A	G	C	A	G	A	G	T	T	T	C	C	A	A	G	A	G	T	A	G	A	G	T	C	T	G	G	T	A	G	A	G	C	A	C	T	G	C	A	G		960		
901	DB	T	T	G	G	T	A	A	A	G	C	A	G	A	G	T	T	T	C	C	A	A	G	A	G	T	A	G	A	G	T	C	T	G	G	T	A	G	A	G	C	A	C	T	G	C	A	G		960		
961	QY	A	G	A	G	C	A	T	C	C	G	C	T	G	T	C	C	G	T	C	T	G	A	G	A	G	C	T	C	T	G	A	G	A	G	C	T	C	C	T	A	C	T	A	G	T	C	C		1020		
961	DB	A	G	A	G	C	A	T	C	C	G	C	T	G	T	C	C	G	T	C	T	G	A	G	A	G	C	T	C	T	G	A	G	A	G	C	T	C	C	T	A	C	T	A	G	T	C	C		1020		
1021	QY	T	T	G	G	T	C	A	C	C	G	T	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C		1080		
1021	DB	T	T	G	G	T	C	A	C	C	G	T	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C	C	A	T	G	C	T	A	C		1080		
1081	QY	T	T	C	T	A	T	A	C	C	A	A	C	A	G	C	A	T	C	T	T	T	G	G	A	A	A	G	C	T	G	G	A	T	C	C	C	T	C	C	G	C	A	A	A	G	A	T	C	C		1140
1081	DB	T	T	C	T	A	T	A	C	C	A	A	C	A	G	C	A	T	C	T	T	T	G	G	A	A	A	G	C	T	G	G	A	T	C	C	C	T	C	C	G	C	A	A	A	G	A	T	C	C		1140
1141	QY	A	C	T	T	G	A	T	A	C	A	G	G	G	C	A	T	C	G	A	C	T	T	T	G	C	C	G	T	T	G	C	C	G	T	T	G	C	C	G	T	T	G	T	C	A	T	T	G		1200	
1141	DB	A	C	T	T	G	A	T	A	C	A	G	G	G	C	A	T	C	G	A	C	T	T	T	G	C	C	G	T	T	G	C	C	G	T	T	G	C	C	G	T	T	G	T	C	A	T	T	G		1200	
1201	QY	C	A	C	T	G	G	A	C	G	A	C	C	C	T	C	C	T	C	A	T	T	G	T	G	C	T	T	T	G	G	C	T	A	T	G	G	C	C	T	C	A	T	T	G	G		1260				
1201	DB	C	A	C	T	G	G	A	C	G	A	C	C	C	T	C	C	T	C	A	T	T	G	T	G	C	T	T	T	G	G	C	T	A	T	G	G	C	C	T	C	A	T	T	G	G		1260				
1261	QY	A	C	C	T	C	A	C	C	T	A	C	C	T	G	A	C	C	A	C	C	T	G	A	C	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		1320			
1261	DB	A	C	C	T	C	A	C	C	T	A	C	C	T	G	A	C	C	A	C	C	T	G	A	C	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		1320			
1321	QY	T	T	G	G	C	A	T	T	C	G	C	C	T	T																																					

Db 841 CGTACCTGCTCTTGGAGAACGCAACGAGCGAAGAGCTGTGAAAGCCTTCCAAACGTTTC 900
QY 901 TTGGGTAAAGCAGACGCTTTCCAGAGGTAGAGAGGTCTGGCTGAGAGCCACGTCGAG 960
Db 901 TTGGGTAAAGCAGACGCTTTCCAGAGGTAGAGAGGTCTGGCTGAGAGCCACGTCGAG 960
QY 961 AGGAGCATCCGCTGGTGTCCGTCGAGCTGCTGAGAGCTCCCTACGTCGCGTCGCGCAG 1020
Db 961 AGGAGCATCCGCTGGTGTCCGTCGAGCTGCTGAGAGCTCCCTACGTCGCGTCGCGCAG 1020
QY 1021 GTGTGTACCGGTGATGTCAACATGGCTGTACACAGCTCTGTGGCTCAATGCAATTTGG 1080
Db 1021 GTGTGTACCGGTGATGTCAACATGGCTGTACACAGCTCTGTGGCTCAATGCAATTTGG 1080
QY 1081 TTCTATACCAACAGCATCTTTGAAAGAGCTGGGATCCCTCCGGCAAGAGATCCCATACGTC 1140
Db 1081 TTCTATACCAACAGCATCTTTGAAAGAGCTGGGATCCCTCCGGCAAGAGATCCCATACGTC 1140
QY 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTGCTTTGGTCAITGAG 1200
Db 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGCTCTTCTGCTTTGGTCAITGAG 1200
QY 1201 CACCTGGGACGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260
Db 1201 CACCTGGGACGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260
QY 1261 ACCCTCACCATCAGCTGACCTGAGGACACAGCCGCTGGGTCCGCTACCTGAGTATC 1320
Db 1261 ACCCTCACCATCAGCTGACCTGAGGACACAGCCGCTGGGTCCGCTACCTGAGTATC 1320
QY 1321 GTGGGCACTTGGCCATCATCGCTCTTTGCGAGTGGCCAGGTGGCATCCGCTTCATC 1380
Db 1321 GTGGGCACTTGGCCATCATCGCTCTTTGCGAGTGGCCAGGTGGCATCCGCTTCATC 1380
QY 1381 TTGACTGTGTGAGTTCTTCCAGCAATCTCAGCGGCGGCTGCTTCATCATTTGCAAGCAC 1440
Db 1381 TTGACTGTGTGAGTTCTTCCAGCAATCTCAGCGGCGGCTGCTTCATCATTTGCAAGCAC 1440
QY 1441 GTCAACTGGCTCTCAACTTTGCTGTGGGCTCTCTTCCATTCATTCAGAAAGTCTG 1500
Db 1441 GTCAACTGGCTCTCAACTTTGCTGTGGGCTCTCTTCCATTCATTCAGAAAGTCTG 1500
QY 1501 GACACTACTGTTCTTCTGCTACAAATTTGCTACAGTGTGATCAGAGTGTATCTACTCTGAT 1560
Db 1501 GACACTACTGTTCTTCTGCTACAAATTTGCTACAGTGTGATCAGAGTGTATCTACTCTGAT 1560
QY 1561 TTGTGTGCTGCTGAGACCAAAACAGAACCTATGCAAGAAATCAGCCAGGCAATTTCCAAA 1620
Db 1561 TTGTGTGCTGCTGAGACCAAAACAGAACCTATGCAAGAAATCAGCCAGGCAATTTCCAAA 1620
QY 1621 AGGAAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGCTGTCACTGATGCTCTGCT 1680
Db 1621 AGGAAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGCTGTCACTGATGCTCTGCT 1680
QY 1681 TCTTCTCTCTTCACTACTCCGAATACAGCTGGAATTCAGCTGCGGCCACCAACCGCC 1740
Db 1681 TCTTCTCTCTTCACTACTCCGAATACAGCTGGAATTCAGCTGCGGCCACCAACCGCC 1740
QY 1741 ACCAAAGAGACACCCATTGTAACGGTCACTGCTGATTTCTCAACTGGAATGACCT 1800
Db 1741 ACCAAAGAGACACCCATTGTAACGGTCACTGCTGATTTCTCAACTGGAATGACCT 1800
QY 1801 TCCCTTATCTTCTCTCTGAGAACACCAAGTCAATGATGTGAGAACAGCTTGGATTT 1860
Db 1801 TCCCTTATCTTCTCTCTGAGAACACCAAGTCAATGATGTGAGAACAGCTTGGATTT 1860
QY 1861 TGGAGACATGGGTTTGAATCCAGTCAATTCATTTTATTCAGCAATATTTAAAGT 1920
Db 1861 TGGAGACATGGGTTTGAATCCAGTCAATTCATTTTATTCAGCAATATTTAAAGT 1920
QY 1921 ACTGACATGCCATATGTTTATCCCACTGGTATACAAATGGAGGGAGAGAGAG 1980
Db 1921 ACTGACATGCCATATGTTTATCCCACTGGTATACAAATGGAGGGAGAGAGAG 1980

QY 1981 AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCAAGTGGCTC 2040
Db 1981 AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCAAGTGGCTC 2040
QY 2041 AGCCCTGTAAATCCAGCACTTTGGAGGCCGAGGTGGGTGGATCGTGAGGTGAGGAGATT 2100
Db 2041 AGCCCTGTAAATCCAGCACTTTGGAGGCCGAGGTGGGTGGATCGTGAGGTGAGGAGATT 2100
QY 2101 GAGACCATCTCTGGCTTAACATGCTGAACTCCCTCTCTACTAAAAATACAAAAATTAGCT 2160
Db 2101 GAGACCATCTCTGGCTTAACATGCTGAACTCCCTCTCTACTAAAAATACAAAAATTAGCT 2160
QY 2161 GAGCATGTGGGGGGCTGTAGTCCAGCTACTTTGGAGGCTGAGCAGGAGATGCG 2220
Db 2161 GAGCATGTGGGGGGCTGTAGTCCAGCTACTTTGGAGGCTGAGCAGGAGATGCG 2220
QY 2221 GTGAACCCAGAGGCGGAGCTTGCAGTGGCGGATCACACCACACTCCAGCCTGG 2280
Db 2221 GTGAACCCAGAGGCGGAGCTTGCAGTGGCGGATCACACCACACTCCAGCCTGG 2280
QY 2281 GTGACAGAGCCAGACTCCGCTCTCAAAAAAAGGCTCTCAAAAAAAGGCGCGC 2340
Db 2281 GTGACAGAGCCAGACTCCGCTCTCAAAAAAAGGCTCTCAAAAAAAGGCGCGC 2340
QY 2341 CGC 2343
Db 2341 CGC 2343

RESULT 3

AAD27918

ID AAD27918 standard; cDNA; 2343 BP.

XX AAD27918;

AC AAD27918;

XX 18-JUN-2002 (first entry)

DT 18-JUN-2002 (first entry)

XX Human glucose transporter protein GLUTX cDNA.

DE Human glucose transporter protein GLUTX cDNA.

XX Glucose transporter; GLUTX; hexose; human; ss.

KW Homo sapiens.

XX Location/Qualifiers

XX 73..1764

FT /tag= a

FT /product= "Glucose transporter, GLUTX"

FT 2337..2339

FT /tag= b

FT /notes= "Nucleotides GGC are present at this location in

FT the sequence shown as SEQ ID NO:1 in Sequence Listing of

FT the specification"

XX US6346374-B1.

PN 12-FEB-2002.

XX 05-JUL-2000; 2000US-00610417.

XX 26-FEB-1998; 98US-00031392.

XX 26-APR-1999; 99US-00299549.

XX (MILL-) MILLENNIUM PHARM INC.

XX Tartaglia LA, Weng X;

XX WPI; 2002-224989/28.

XX P-PSDB; AAE14611.

XX Identifying compound modulating activity of glucose transporter, useful

XX for treating disorders mediated by transporter, comprises contacting

XX polypeptide with test compound and comparing activity in presence and

CC also useful as a therapeutic agent for regulating translation of GLUTX
CC mRNA, and for treatment of disorders associated with aberrant expression
CC of GLUTX and aberrant hexose transport. It is useful for generating GLUTX
CC specific antibodies, identifying agonists and antagonists of GLUTX, and
CC identifying nucleic acids in other species encoding nucleic acids
CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying
CC the chromosomal location of GLUTX, and as tissue specific markers
XX
SQ Sequence 2343 BP; 535 A; 637 C; 621 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 2343; DB 2; Length 2343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGACCCACAGCTCCGGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60
Db	1	TCGACCCACAGCTCCGGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60
Qy	61	TCACTGAGACCCATGGCAAGAAACAAATAGGAATCCAAAGAACTGGGCCCTAGTTC	120
Db	61	TCACTGAGACCCATGGCAAGAAACAAATAGGAATCCAAAGAACTGGGCCCTAGTTC	120
Qy	121	CTCAGATGACACAGCAGCCCGGCTCCAGGGCCAGGAGGACCTGCTGGAGTGT	180
Db	121	CTCAGATGACACAGCAGCCCGGCTCCAGGGCCAGGAGGACCTGCTGGAGTGT	180
Qy	181	GACCACTGAGAGTGGGTCCAGGTGGAGGAGAAAGACCTGCTCTGCTC	240
Db	181	GACCACTGAGAGTGGGTCCAGGTGGAGGAGAAAGACCTGCTCTGCTC	240
Qy	241	CTGTGGCTCCCTCGGGGGCCCTTCGGTCTCTCTTCTACGGCTACAACTGTG	300
Db	241	CTGTGGCTCCCTCGGGGGCCCTTCGGTCTCTCTTCTACGGCTACAACTGTG	300
Qy	301	GTGTGTAATGCCCAACCCCTAGATCAAGGCCCTTTTACATGAGTCAATGGAAAGG	360
Db	301	GTGTGTAATGCCCAACCCCTAGATCAAGGCCCTTTTACATGAGTCAATGGAAAGG	360
Qy	361	CATGGAAGTCCAAATAGACCCAGACCTGACTCTGCTCTGCTGCTGACTGTCCATA	420
Db	361	CATGGAAGTCCAAATAGACCCAGACCTGACTCTGCTCTGCTGCTGACTGTCCATA	420
Qy	421	TTGCGCATGCTGACTTTGGGACGTTAATGTGAAGATGATTGGAAGGTTCTTGGG	480
Db	421	TTGCGCATGCTGACTTTGGGACGTTAATGTGAAGATGATTGGAAGGTTCTTGGG	480
Qy	481	AGGAAGCACACTTTGCTGGCCAAATATGGGTTCGCAATTTCTGCTGATTCGATGCC	540
Db	481	AGGAAGCACACTTTGCTGGCCAAATATGGGTTCGCAATTTCTGCTGATTCGATGCC	540
Qy	541	TGCTGCTCCAGGAGGAGCCCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGG	600
Db	541	TGCTGCTCCAGGAGGAGCCCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGG	600
Qy	601	GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG	660
Db	601	GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG	660
Qy	661	ATCCGTGACTCTCGGGGACGTGACTGCCATCTTTATCTGCTGCTGCTGCTGCTGCTG	720
Db	661	ATCCGTGACTCTCGGGGACGTGACTGCCATCTTTATCTGCTGCTGCTGCTGCTGCTG	720
Qy	721	CAGCTTCTGGGCTGCTGCGGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGA	780
Db	721	CAGCTTCTGGGCTGCTGCGGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGA	780
Qy	781	GTGATGTGGTCCCTGCGGTTGTCAGCTGTGAGGCTTCCCTTCTCCCGGACAGCCCA	840
Db	781	GTGATGTGGTCCCTGCGGTTGTCAGCTGTGAGGCTTCCCTTCTCCCGGACAGCCCA	840
Qy	841	CGCTACCTGCTCTGGAGAGCAACACAGGACAGAGCTGTGAAGCTTCCAAACGTTT	900
Db	841	CGCTACCTGCTCTGGAGAGCAACACAGGACAGAGCTGTGAAGCTTCCAAACGTTT	900

Qy	901	TTGGTAAAGCAGAGCTTTCCAGAGGTAGAGGAGTCTTGGCTGAGAGCACGTCGAG	960
Db	901	TTGGTAAAGCAGAGCTTTCCAGAGGTAGAGGAGTCTTGGCTGAGAGCACGTCGAG	960
Qy	961	AGGAGCATCCGCTCGTCTCGTCTGAGTCTGAGAGTCCCTACGTCGCTGGGAG	1020
Db	961	AGGAGCATCCGCTCGTCTCGTCTGAGTCTGAGAGTCCCTACGTCGCTGGGAG	1020
Qy	1021	GTGGTCAACGTAATGTCCACATGGCCCTGACAGCTCTGTGGCTCAATGCAATTTGG	1080
Db	1021	GTGGTCAACGTAATGTCCACATGGCCCTGACAGCTCTGTGGCTCAATGCAATTTGG	1080
Qy	1081	TTCTATATCAACAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATAC	1140
Db	1081	TTCTATATCAACAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATAC	1140
Qy	1141	ACCTTGAATACAGGGGAGCATCGAGACTTTGGTGGCTCTCTCTGCTGCTGCTG	1200
Db	1141	ACCTTGAATACAGGGGAGCATCGAGACTTTGGTGGCTCTCTCTGCTGCTGCTG	1200
Qy	1201	CACCTGGGAGGAGACCTCTCTCTATTTGGTCTTGGGCTCATGGGCTCTCTCTTGG	1260
Db	1201	CACCTGGGAGGAGACCTCTCTCTATTTGGTCTTGGGCTCATGGGCTCTCTCTTGG	1260
Qy	1261	ACCTTCAACATCACGCTGACCTGACAGCAACGCCCCCTGGTCCCTACTCTGATATC	1320
Db	1261	ACCTTCAACATCACGCTGACCTGACAGCAACGCCCCCTGGTCCCTACTCTGATATC	1320
Qy	1321	GTGGGCAATTTGGCCATCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
Db	1321	GTGGGCAATTTGGCCATCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
Qy	1381	TTGACTGTGAGTCTCTCCAGCAATCTCAGCGGCGGCTTCTCATCTGAGGACAC	1440
Db	1381	TTGACTGTGAGTCTCTCCAGCAATCTCAGCGGCGGCTTCTCATCTGAGGACAC	1440
Qy	1441	GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCT	1500
Db	1441	GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCT	1500
Qy	1501	GACACTACTGTTTCTTCTAGTCTTTGCTCAATTTGATCAGAGTGTCTATCTAC	1560
Db	1501	GACACTACTGTTTCTTCTAGTCTTTGCTCAATTTGATCAGAGTGTCTATCTAC	1560
Qy	1561	TTTGTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCC	1620
Db	1561	TTTGTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCC	1620
Qy	1621	AGGACAAAGCATATCCACAGAGAAATTCGACTCAGCTGTCTCACTGATCTCTCT	1680
Db	1621	AGGACAAAGCATATCCACAGAGAAATTCGACTCAGCTGTCTCACTGATCTCTCT	1680
Qy	1681	TCTTCTCTCTTCTCAGTCTCCGAAATCAGCTGGATTCAGCTGCGCCACACCG	1740
Db	1681	TCTTCTCTCTTCTCAGTCTCCGAAATCAGCTGGATTCAGCTGCGCCACACCG	1740
Qy	1741	ACCAAAAGAACACCCATTTGAAAGGCTATGTTTCTCACTGATCTGATCTGAT	1800
Db	1741	ACCAAAAGAACACCCATTTGAAAGGCTATGTTTCTCACTGATCTGATCTGAT	1800
Qy	1801	TCCCTATCTCTTCTCTGAGAACACCAAGTCTGATGTCAGACAGAGCTTGGAT	1860
Db	1801	TCCCTATCTCTTCTCTGAGAACACCAAGTCTGATGTCAGACAGAGCTTGGAT	1860
Qy	1861	TGGAGACATGGGTTTGAATTCAGTCAATTTTATTTCAGCAAAATATTAAAGT	1920
Db	1861	TGGAGACATGGGTTTGAATTCAGTCAATTTTATTTCAGCAAAATATTAAAGT	1920
Qy	1921	ACTGACATCTCCCATATGTTTGTATACCACTGTTTATACCACTGTTTATACCA	1980
Db	1921	ACTGACATCTCCCATATGTTTGTATACCACTGTTTATACCACTGTTTATACCA	1980

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 21:02:47 ; Search time 922 Seconds
(without alignments)
10795.584 Million cell updates/sec

Title: US-09-981-947B-1

Perfect score: 2343

Sequence: 1 tcgaccacgcgtccggcct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2343	100.0	2343	2	Aaz11502 Human glu
2	2343	100.0	2343	5	Aac62143 CDNA enco
3	2343	100.0	2343	6	Aad27918 Human glu
4	1651	70.5	2032	6	Aad27278 Human tra
5	1644.6	70.2	1851	4	Abn09154 Human GLU
6	1644.6	70.2	1851	4	Aak52555 Human pol
7	1643	70.1	2242	4	Aai60472 Human pol
8	1399	59.7	1710	9	Adc26250 Human NOV
9	1317.8	56.2	1704	4	Aak51571 Human pol
10	1317.8	56.2	2842	4	Aai58686 Novel hum
11	1317.8	56.2	2842	8	Adb48664 Novel hum
12	1232.8	52.6	1757	9	Adc26252 Human NOV
13	1229	52.5	1502	9	Adc26254 Human NOV
14	1187.6	50.7	1267	9	Adc26264 Human NOV
15	1141	48.7	2216	4	Aad09557 Human tra
16	1133.8	48.4	1189	9	Adc26260 Human NOV
17	1130.6	48.3	1189	9	Adc26258 Human NOV
18	1130.6	48.3	1189	9	Adc26262 Human NOV
19	915	39.1	1087	9	Adc26256 Human NOV
20	430.8	18.4	472	8	Ach43430 Human foe
21	427.6	18.3	1536	9	Adn01366 Human TCH
22	427.6	18.3	1539	9	Acc42990 Human glu
23	427.6	18.3	1563	9	Adn01388 Human var

24	427.6	18.3	1552	9	ADD01378 Human TCH
25	424.4	18.1	1539	6	Aal44673 Human tra
26	421.6	18.0	2169	7	Abt42242 Toxicity
27	421.6	18.0	2169	9	Adb53334 Primary r
28	416.4	17.8	2218	6	Abi64747 Lung canc
29	416.4	17.8	2218	6	Abi65386 Lung canc
30	416.4	17.8	2218	6	Abv78029 Hypoxia-r
31	416.4	17.8	2218	7	Acc72731 Human can
32	410	17.5	7434	6	Adb28386 Human che
33	407.8	17.4	7434	6	Adb28387 Human che
34	376	16.0	1607	6	Abss52105 Human glu
35	354	15.1	1461	7	Acc05717 CDNA enco
36	354	15.1	2064	7	Aad53978 Human cdn
37	346	14.9	2222	5	Aaf93832 Human sbg
38	335.2	14.3	1491	6	Abi57737 Human sbg
39	334.8	14.3	1501	6	Aal44681 Human tra
40	333.2	14.2	2135	6	Abt31931 Human bre
41	330.2	14.1	2778	7	Abx74453 Human cdn
42	330.2	14.1	2854	7	Ada28225 Human lun
43	330.2	14.1	2856	8	Ada28225 Human lun
44	330.2	14.1	2893	5	Aaf98714 Human lac
45	330.2	14.1	3366	6	Abz35362 Human gen

ALIGNMENTS

RESULT 1

Aaz11502

ID AAZ11502 standard; cDNA; 2343 BP.

XX

AC AAZ11502;

XX

DT 05-NOV-1999 (first entry)

XX

DE Human glucose transporter protein GLUTX encoding cDNA.

XX

KW Human; glucose transporter; GLUTX; hexose transport; gene therapy;

KW Human; glucose transporter; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS 73..1764

FT /*tag= a

FT /product= "GLUTX"

XX

PN US942398-A.

XX

PD 24-AUG-1999.

XX

PF 26-FEB-1998; 98US-00031392.

XX

PR 26-FEB-1998; 98US-00031392.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Tartaglia La, Weng X;

XX

DR WPI; 1999-526192/44.

XX

P-PSDB; AAY27286.

XX

PT A novel mammalian transporter protein and the gene encoding it, are

XX

PT useful in the diagnosis and treatment of disorders associated with

XX

PT aberrant sugar transport.

XX

PS Claim 2; Fig 1A-E; 48pp; English.

XX

CC This cDNA encodes a human glucose transporter (GLUTX) protein. The

CC protein can be expressed by standard recombinant methodology. GLUTX

CC nucleic acids are useful as hybridization probes for detecting the

CC presence of GLUTX DNA in a sample, useful for diagnosing conditions

CC associated with aberrant expression levels of GLUTX. The GLUTX gene is

Qy 956 TGCAGAGGAGCATCCGCTCGGTGTCGGTGTGGAGCTGTGAGAGCTCCCTAGTCCGCT 1015
Db 785 CCGAGCGCGGAGGCGCACCTGTCTGTGTGACCTCTGTGCCCTCGGTCCTCGCT 844
Qy 1016 GGCAGGTGGTCACCGTGAATGTACCATGGCTTGTACAGCTCTGTGGCTCAATGCAA 1075
Db 845 GGCAGCTCCTTCCTCATCTGTCTCATGTGCCCGCCAGAGCTGTGGGCAATCAATGCGA 904
Qy 1076 TTTGGTTCTTATACCAACAGCATCTTTGAAAAAGTGGATCCCTCCGGGCAAGATCCCAT 1135
Db 905 TCAACTACTATGCGGACACCATCTACACATCTGGGGCGTGGAGGCGCTCACTCCCAAT 964
Qy 1136 ACGTCACTTTAGTACAGGGGCGCATCGAGACTTTGGCTGCCGTCTTCTCTGTGGTTGGTCA 1195
Db 965 ATGTAACGGTGGGCTCTGGGCTCGTCAACATAGTAGATGACATCACCTCGCTGTCTTG 1024
Qy 1196 TTGAGCACTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCT 1255
Db 1025 TGGAGCGGCTGGGACGCGGCGACCTCTCTGTGGCGGCTACGGCACTCGGGCTCTGCCT 1084
Qy 1256 TTGGGACCTCACCATCACGCTGACCTTGAGGACACAGCCCTCGGGTCCCTTACCTGA 1315
Db 1085 GCCTGGTGTACGGTGGTGTCTCTATTCCAGAACAGGGTCCCGAGCTGTCTTACCTCG 1144
Qy 1316 GTATCGTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCCAGGTGGCATCCCGT 1375
Db 1145 GCATCATCTGTCTTTGGCTTACATCGGGGACATTCATTTGGGCCAGTCTCTGTCCCT 1204
Qy 1376 TCATCTTCACTGGTGAATTTCCAGAAATCTCAGCGCGGCTGCTTCTCATATTGCAG 1435
Db 1205 CGGTGGTGAAGACCGAGATCTTCTGCACTCTCCCGCGGGCAGCTTTCATGTTGACG 1264
Qy 1436 GCACGTCACACTGGCTCTCCAACTTGTCTGTTGGCTCTCTTCCCATTCATTCAGAAAA 1495
Db 1265 GGGCAGTGCCTGGGCTCAACCACTTCATATAGGCTTCCCTGTCCCATCCATCCAGAGG 1324
Qy 1496 GTCTGGACACCTACTGTTTCTAGTCTTTGTGTACATTTGTATCACAGGTGCTATCTACC 1555
Db 1325 CCATCGGTGCCCTACAGTTTTCATCTTTTGGGAACTGCTCTCTCACTGGGATTACA 1384
Qy 1556 TGTATTTTGTGCTGCTGAGACCAAAAAACAGAACCTATGCAGAAATCAGCCAGGCATTTT 1615
Db 1385 TCTACGTGGTTATTCGGAGACCAAGGCAAAACATTTGTGGAGATAACCGCATTTTGT 1444
Qy 1616 CCAAGAGGAACAAGCATACCCACAGAGAGAAATCGACTCAGCTGTCTACTGTGCTC 1675
Db 1445 CCAAGAGAAACAGGGTGAAGCTTCCAGAGGAGAAAGAAACCATTTGATGCTGGGCTC 1504
Qy 1676 CTGCTTCTCTCTTCA 1693
Db 1505 CCACAGCCTCTCTGCCA 1522

Search completed: April 8, 2004, 14:27:32
Job time : 9037 secs

AX753091	AX753091	AX753091	1539 bp	DNA	linear	PAT 23-JUN-2003
LOCUS	Sequence 1 from Patent WO03038094.					
DEFINITION	AX753091					
ACCESSION	AX753091					
VERSION	AX753091.1	GI:32165845				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1. Xiac, Y.					
TITLE	Regulation of human glucose transporter					
JOURNAL	Patent: WO 03038094-A 1 08-MAY-2003;					
FEATURES	Bayer Aktiengesellschaft (DE)					
source	Location/Qualifiers					
	1. .1539					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
	Query Match	18.3%	Score 427.6;	DB 6;	Length 1539;	
	Best Local Similarity	55.8%;	Pred. No. 1.1e-91;			
	Matches 814;	Conservative 0;	Mismatches 644;	Indels 0;	Gaps 0;	
QY	236	CGCTCCTCGTGGCTCCCTCGCGGGCGCTTCGGCTCTCCTTCTCTCTACGGCTACAC	295			
DB	65	CGCTGTGTGGCGACACTGAGCGGGCTTTGGCTGAGCCTTCAGTACGGCTACAC	124			
QY	296	TGTCGGTGTCAATGCCCCACCCCGTACATCAAGGCCCTTTTCAATGAGTCATGGGAA	355			
DB	125	TCTCTGTGGTCAACACGCCGCAAGGCTTTCAGTCAATTTACAAAGAACTTACTTG	184			
QY	356	GAAGCATGAGACGTCAATAGACCCAGACACTCTGACTCTGCTCTGGTCTGTGACTGT	415			
DB	185	AGCACGCAACCATTCATGGAACGGAAGCTCATGCTGCTTCTATGGCTTGCACCGT	244			
QY	416	CCATATTGCGCATCGGTGACTGTGGGACGCTTAATTGTGAAGATGATTGGAAGGTT	475			
DB	245	CCATGTTTCTCTGGCGGCTGTGGGTCATGCTCTGGGCTGCTGGTGCATAGCT	304			
QY	476	TGGAGAGAAAGACACATTTGCTGGCCAAATATGSGTTTGCATTTCTGTGCATGTCTGA	535			
DB	305	CGCGAGAAAGGGGACCTGCTGATCAACAAATCTTTGCCATCATCCCCGCATCTCTGA	364			
QY	536	TGGCCTGCTCGCTCCAGGACGAGGCTTTGAAATGCTCATTTGTGGGAGCTTATCATG	595			
DB	365	TGGAGTCACAAAGTGGCCAAAGCTTTTGAAGTCATCGTCTTTCCGAGTGTGTCTGG	424			
QY	596	GCATAGATGAGGCGTTCGCCCTCAGTGTGCTCCCATGATACCTCAGTGAATCTCAC	655			
DB	425	GAGTCTGTGAGCATCTCCTACAGCGCCCTCCCATGATCCTGGGAGAACTGGCCCC	484			
QY	656	AGGATGCTGTGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGATGGGGTGTCTCA	715			
DB	485	AGAACTTGAGAGGCATGGTGGGAAACAATGAACGAGGTTTTCGTATCGTTGGAGTCT	544			
QY	716	CTGGGACGCTCTGGGCTGCCCGAGCTGCTGGGAAAGGAGTACCTGGCCATACCTGT	775			
DB	545	TAGCACAGATCTTACGCTCCAGGCCAATCTTGGGCAACCCGGAGGCTGGCGGTCTTC	604			
QY	776	TGGAGTGTATGTGTGCTCTGCGGTGTGTCAGCTGTCTGAGCCTTCCCTTTCTCCCG	835			
DB	605	TGGCGCTCACAGGGGTGCCCGCCTGCTGACGTGCTGACCTGCCCTTCTTCCCGGAA	664			
QY	836	GCCACGCTACCTGCTCTTGGAGAGACACACGAGGCAAGAGCTGTGAAGCCTTCCAA	895			
DB	665	GCCCCGCTACTCCCTGATTCAGAAAGAGATGAAGCCACAGCCGACAGCTCTGAGGA	724			
QY	896	CGTCTTGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAGGTCTCGCTGAGAGCCAG	955			
DB	725	GGCTGAGAGGCCACACGACATGAGGCGGAGCTGGAGGACATGCTGCGAGAGCCCGG	784			

Db 725 GGCTGAGAGGCCACCGGACATGGAGGCCGAGCTGGAGGACATGCGTGGGAGGCCCGG 784

QY 864 CAACGAGCAAGAGCTGTGAAGCCCTCCAAACGCTTCTTGGTAAAGCAGCGTTTC 920
 Db 541 CAACGAGCAAGAGCTGTGAAGCCCTTCCAAACGCTTCTTGGTAAAGCAGCGTTTC 597

RESULT 13
 AF490463 1296 bp mRNA linear ROD 28-MAY-2003
 DEFINITION Mus musculus GLUT9a mRNA, complete cds.
 ACCESSION AF490463
 VERSION AF490463.1 GI:31088266
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Carayannopoulos, M.O., Schlehn, A.L., Hoeft, A.L., Moley, J.F. and
 Moley, K.H.
 Cloning and functional analysis of mouse GLUT9, a new embryonically
 expressed transporter
 Unpublished
 2 (bases 1 to 1296)
 Carayannopoulos, M.O. and Moley, K.H.
 Direct Submission
 Submitted (06-MAR-2002) OB/Gyn, Washington University, 4911 Barnes
 Hospital Plaza, St. Louis, MO 63110, USA
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..1296
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /chromosome="5"
 /dev_stage="embryonic day 7"
 1..1296
 /note="member of the facilitative glucose transporter"
 /codon_start=1
 /product="GLUT9a"
 /protein_id="AAP44162.1"
 /db_xref="GI:31088267"
 /translation="MDSRELALASLMDTGGPGLSVGHQRTKWSFSLVVAALVG
 AFGSFLYGNLSVNVNAPYKAFYNGTVRRHQPDIDPTLTLWSVTVSFAITGG
 LVGTLMKRMFLGRKSTLLVNGFAISALLMACSLRAGTPEMLIVGVFIMGVGG
 IALSALPMVLEISPEIKRSLQGVTAIFICIGVSGQLLGLPELLGRIFWYNSIFG
 KAGIPDKIPYITLSTGGIETLAFISGLVIERLGRPLLGSGFLMALPFGTLTATL
 TLQDQAPVPLISVICILALIASPCSPGGIPFLITGEFQQSERPAFNIAGTVNWL
 SNFVGLLFPFFQKLSYCFILVFATICAGATIFYFVLPEITKNRTHAEISQAFARN
 KACPPEVKAASAMTEKANSOTEPDSSSTLDSYGNKIV"
 627
 /note="deletion of transmembrane regions 6 and 7"

misc_feature
 22.3%; Score 523.6; DB 10; Length 1296;
 Best local Similarity 65.8%; Pred. No. 9.6e-115;
 Matches 962; Conservative 0; Mismatches 179; Indels 321; Gaps 1;

ORIGIN
 211 AGGAGAGAAAGACTGGTCCGCTCGCTCGTGGGCTCCCTCGGGGGCGCTTCGGC 270
 Db 82 AGGAGACAAAGAAATGGTCTCTCGCTCGTGGGAGCTTCGTTGGGAGCTTCGG 141
 QY 271 TCCTCTCTCTCTAGCTTACAACTGCTGGTGGTGAATGCCCAACCCGCTACATCAAG 330
 Db 142 TCCTCTCTCTCTAGCTTACAACTGCTGGTGGTGAATGCCCAACCCGCTATATCAAG 201
 QY 331 GCCTTTTCAATGAGTATGGGAAAGAGGATGAGCTCCCAATAGACCAGACACTCTG 390
 Db 202 GCCTTTTCAATGAGTATGGGAAAGAGGATGAGCTCCCAATAGACCAGACACTCTG 261
 QY 391 ACTCTGCTCTGCTGACTGTGCTTCCATATTCCGCTCGGTGACTTGTGGGAGCTTA 450
 Db 262 ACTCTGCTCTGCTGACTGTGCTTCCATATTGCGATTGCTGGCTGGTGGGAAAGCTG 321

QY 451 ATTGTGAGATGATTGGAAAGTTCTTGGGAGGAGCAGACATTTGCTGGCCATATATGG 510
 Db 322 ATGTGTAAGATGATTGGAAGTTCTTGGGAGGAGTCCACATTTGCTGGTCAACATGGA 381
 QY 511 TTTCAATTTCTGCTGATTTGCTGATGGCTGCTCGCTCCAGCAGAGAGCTTTGAAATG 570
 Db 382 TTTGCCATTTCCGAGCGTTGCTGATGGCATGCTCACTCCGAGCAGAACTTTGAGATG 441
 QY 571 CTCATTGTTGGAGCTTTCATCATGGCATAGATGAGGAGCTCGCCCTCAGTGTGCTCCC 630
 Db 442 CTCATTGTTGGAGCTTTCATCATGGCATAGGATGAGGAGCTTGCATCTCAGCGCTCCC 501
 QY 631 ATGTACCTCAGTGAATCTCACCAAGGAGATCCGTGGCTCTCTCTGGGGCAGGTGACTGCC 690
 Db 502 ATGTATCTCAGGAGATCTCACCAAGGAGATCCGGGGCTCTCTGGGGCAGGTGACTGCC 561
 QY 691 ATCTTTATCTCATTTGGCTGTTCACTGGGAGCTTCTGGGCTGCTGGCCAGAGCTGGGA 750
 Db 562 ATCTTCATCTCATTTGGCTGTTTCTTGGACAGTTGCTGGGCTGCTGGGAGCTGGGA 621
 QY 751 AAGGAGAGTACCTGGCCATACCTGTTGGAGTATTGCTGCTCCCTGCTGCTCCAGCTG 810
 Db 622 AGG----- 624
 QY 811 CTGAGCCTTCCCTTTCTCCGGAGCAGCCACGCTACCTGCTCTTGGAGAGCAGACAGAG 870
 Db 625 ----- 624
 QY 871 GCAAGAGCTGTGAAGAGCTTCCAAAAGTCTTGGGTAAAGCAGACGTTTCCCAGAGGTA 930
 Db 625 ----- 624
 QY 931 GAGGAGGCTCTGGCTGAGAGCCAGCTGACAGGAGCATCCGCTGCTGCTCGTGGAG 990
 Db 625 ----- 624
 QY 991 CTGCTGAGAGCTCCCTAGCTCCGCTGAGGAGTGGTCAACGCTGATGCTCACCATTGGCTGC 1050
 Db 625 ----- 624
 QY 1051 TACAGCTCTGTGGCTCAATGCAATTTGGTTCTTACCAACAGCATCTTTGGAAGCT 1110
 Db 625 ----- ATCTGTTTCTACCAACAGCATCTTTGGAAGCT 660
 QY 1111 GGGATCCCTCGGCAAGATCCCATACGTCACCTGAGTACAGGGGGCATTCAGACTTGG 1170
 Db 661 GGGATCCCTCAGGACAAGATCCCATACATCCCTGAGCAGCGAGGAATTGAAACACTG 720
 QY 1171 GCTGCGCTTCTCTGTTTGGTCAATGAGCAGCTGGGAGAGACCCCTCCTCATTTGT 1230
 Db 721 GCTGCTATCTCTCTGCTTGGTCAATGAGCGCTTAGGAAGGAGACCTCTCCTCATTTGT 780
 QY 1231 GGGTTTGGGCTCATGGGCTCTTCTTTGGGAGCCCTCACCATCAGCTGACCTCGAGGAG 1290
 Db 781 GGGTTTGGGCTCATGGGCTCTTCTTTGGGAGCCCTCCTTGGAGACCTCAGCGACACTGACCTCGAGGAC 840
 QY 1291 CACGCCCCCTGGTCCCTTACCTGAGTATCGTGGGATTTGCGCCATCATGCGCTTTTC 1350
 Db 841 CAACTCCCTGGGCTCCCTTACCTCAGCATTTGTCATCTTGGCCATCATTTGCTGCTTC 900
 QY 1351 TGGAGTGGGAGAGTGGGATCCCGTTCATCTTCACTGGTGGTCTTCTCCAGCAATCTCAG 1410
 Db 901 TGCAGCGCTCCAGGTGGGATCCCATCTCTACCCGAGAGTCTTCCAAAGTCAAGAG 960
 QY 1411 CGGCGGCTGCTCTCATCTATTGAGGAGCAGCTCACTGGCTCTTCCAACTTTGCTGTGGG 1470
 Db 961 CGACCGGAGCGCTTCAATGATCGAGGAGCAGTCACTGGCTCTCAAACTTCGCGGTAGGG 1020
 QY 1471 CTCCTCTTCCCATTCATTCAGAAAAGTCTGGACCACTACTGTTTCTCTAGTCTTTGCTACA 1530
 Db 1021 CTCCTTTTCCCTTTCATCCAGAAAAGCTGGACTCTCTACTGTTCTCTCTCTCTGCGACA 1080
 QY 1531 ATTGTATCAGAGTGTCTATCTACCTGATTTTGTGCTGCTGAGACCAAAACAGAAC 1590

Db 111121 GTCAGGAGATTGAGACCACTCTGGCTACATGGTGAAACTCCCTCTCTACTTAAATAACA 111062
 QY 2150 AAAATTAGCTGAGCATGGTGGCGGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGC 2209
 Db 111061 AAAATTAGCTGAGCATGGTGGCGGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGC 111002
 QY 2210 AGGAGATGGCGTGAACCCAGGAGCGGAGTTGCGAGTGAGCCGAGATCACACACACA 2269
 Db 111001 AGGAGATGGCGTGAACCCAGGAGCGGAGTTGCGAGTGAGCCGAGATCACACACACA 110942
 QY 2270 CTCAGCTGGTGTGACAGGCGAGCTCCGTCTCAAAAAAATAAAAAA 2326
 Db 110941 CTCAGCTGGTGTGACAGGCGAGCTCCGTCTCAAAAAAATAAAAAA 110885

RESULT 10
 AC098976/c
 LOCUS Homo sapiens BAC clone RP11-751L19 from 4, complete sequence.
 DEFINITION AC098976 AC068403
 ACCESSION AC098976.2 GI:18042509
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Sulston, J.B. and Waterston, R.
 1 (bases 1 to 165221)
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99063792
 9847074

REFERENCE 2 (bases 1 to 165221)
 AUTHORS Levy, A. and Cotton, M.
 TITLE The sequence of Homo sapiens BAC clone RP11-751L19
 JOURNAL Unpublished (2001)
 MEDLINE
 3 (bases 1 to 165221)
 REFERENCE 3 (bases 1 to 165221)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (07-NOV-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 165221)
 REFERENCE 4 (bases 1 to 165221)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 165221)
 REFERENCE 5 (bases 1 to 165221)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jan 3, 2002 this sequence version replaced gi:16756380.

COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 Center project name: H_NH0751L19
 Drafting Center: WIBR

 NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.
 This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.
 MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,
 Tatenio, M., Catanese, J.J., and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP11-751L19;
 actual end is at base position 165221 of RP11-751L19.

Data from AC067775 was used to finish this clone, AC098976.
 The sequence of AC068403 has been incorporated into AC098976.

FEATURES	Location/Qualifiers
source	1..165221
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-751L19"
	/clone_lib="RPCI-11"
misc_feature	1..154
	/note="match to EST B5062468 (NID:G8407118)"
misc_feature	161..427
	/note="similar to Homo sapiens EST BF878566 (NID:G12368696)"
repeat_region	173..290
	/rpt_family="L2"
repeat_region	915..1201
	/rpt_family="Alu"
repeat_region	1203..1425
	/rpt_family="MER2_type"
repeat_region	1638..1952
	/rpt_family="Alu"
repeat_region	1823..1954
	/rpt_family="(A)n"
repeat_region	3476..3772
	/rpt_family="Alu"
repeat_region	3773..3909
	/rpt_family="ERV1"
misc_feature	3788..3793
	/note="similar to Homo sapiens EST BF171573 (NID:G13437787)"
repeat_region	3910..4193
	/rpt_family="MaLR"
repeat_region	4194..4507
	/rpt_family="Alu"
misc_feature	4194..4208
	/note="similar to Homo sapiens EST AW451324 (NID:G6992100)"
misc_feature	4194..4206
	/note="similar to Homo sapiens EST AI371376 (NID:G4150129) G083f11.xl"
misc_feature	4194..4206
	/note="similar to Homo sapiens EST AI740932 (NID:G5109220) wgl8e02.xl"
repeat_region	4508..4704
	/rpt_family="MaLR"

1161	GGGATCCCTCAGGACAAGATCCCATACATCACCTGTAGACACGGGAGGAATTGAAACACTG	1222
1171	GCTGCGGCTTCTCTGGTTGGTGCATATGAGACACTGGGACGGAGACCCCTCCCTCATTTGGT	1230
1221	GCTGCTATCTTCTCTGGCTTGGTGCATATGAGCGCTTAGGAAGGGGACCTCTCCTCATTTGGT	1280
1231	GGCTTTGGGCTCATGGGCGCTCTTCTTTTGGGACCCCTCACCATCACGCTGACCCCTGCAGGAC	1290
1281	GGCTTTGGGCGCTCATGGGCGCTCTTCTTTTGGGACCCCTCACCGGACACTGACCCCTGCAGGAC	1340
1291	CACGCGCCCTCGGATCCCTACCTGACTGAGTATCGTGGGCATTCTGGCCCATCATCGCCTCTTTC	1350
1341	CAAGCTCCCTGGGTCCCTTACTCTAGCATGTGTGCATCTTGGCCATCATTTGCTCTGTTTC	1400
1351	TGCAGTGGGCCAGGTGGCATCCCGTTTCATCTTTGACTGTGTGAGTTCTTCCAGCAATCTCAG	1410
1401	TGCAGCGGTCCAGGTGGGATCCCATTCATCTTGACGGAGAGTTCTTCCAAACAGTCAGAG	1460
1411	CGGCGGCTGCCTTCATCATTTGCAGGACCCGTCAACTGSGCTCTCMACTTTGCTGTTGGG	1470
1461	CGACGGCGAGCCTTCATGATCGCAGGACAGTCAACTGGCTCTCAAACCTTGCOCGTAGGG	1520
1471	CTTCCTTTCCCATTCATTACGAAAGTCTGGACACTACTGTTTCTTCTAGTCTTTGTGTACA	1530
1521	CTTCCTTTTCCCTTCATCCAGAAAAGCCTGGACTCCCTACTGTTCTCTCGTCTTCGCCACA	1580
1531	ATTTGTATCACAGGTGCTATCTACCTGTATTTTGTCTGCTCGCTGAGACCAAAAACAGAAAC	1590
1581	ATATGTATCGAAGGTGCTACCTACTCTCTATTTTGTCTCCCGGAGACCAAGACAGAAAC	1640
1591	TATGCAGAAATCAGCGCGCATTTTCCAAAAGGAAACAAAGCATACCCACGAGAGAGAAA	1650
1641	CATGCAGAAATCAGCGCGCATTTTGCAGAGAGAAACAAAGGCTCAGGCCCCCGAGAGTGAAG	1700
1651	ATCGACTCGCTGCTCACTGATG	1672
1701	GCTGATTCAGCCATGACTGAGG	1722

RESULT 9	AC067775	150402 bp	DNA	linear	HTG 24-AUG-2002
LOCUS	AC067775/c				
DEFINITION	Homo sapiens chromosome 4 clone RP11-690D17 map 4, WORKING DRAFT SEQUENCE, 9 unordered pieces.				
ACCESSION	AC067775				
VERSION	AC067775.2	GI:8318589			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				

REFERENCE
1 (bases 1 to 160402)
BIRREN, B., LINTON, L., NUSBAUM, C. and LANDER, E.
AUTHORS
TITLE
Homo sapiens chromosome 4, clone RP11-690D17
1 (bases 1 to 160402)

Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 160402)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

Boguslavskiy, L., Bouktgalter, B., Brown, A., Burkett, G.,
Campaniano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,

Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,

Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,

Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murray, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,

O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisaní, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

Roy, A., Santos, R., Severy, S., Severy, P., Spencer, B.,

Db 1657 TCAGCTGTCACTGATGGT 1674
 |||||
 RESULT 6
 AR339083
 LOCUS AR339083 2842 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 574 from patent US 6569662.
 ACCSSION AR339083
 VERSION AR339083.1 GI:33725940
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2842)
 AUTHORS Tang,Y.T., Zhou,P. and Drmanac,R.T.
 TITLE Nucleic acids and polypeptides
 JOURNAL Patent: US 6569662-A 574: 27-MAY-2003;
 FEATURES
 Location/Qualifiers
 1..2842
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 56.2%; Score 1317.8; DB 6; Length 2842;
 Best Local Similarity 89.9%; Pred. No. 1.9e-305;
 Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 3;

QY 2 CGACCCACGCGTCCGGCGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCATCAGCTGGGT 61
 Db |||||
 QY 40 CGACCCACGCGTGC-GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCATCAGCTGGGT 98
 Db |||||
 QY 62 CACTGAGACCCATGGCAAGGAAACAAATATAGGAATTCGAAGAACTGGGCTAGTTCCCG 121
 Db |||||
 QY 99 CACTGAGACCCATGGCAAGGAAACAAATATAGGAATTCGAAGAACTGGGCTAGTTCCCG 158
 Db |||||
 QY 122 TCACAGATGACACGAGCCACCGCGGCTCCAGGGCCAGGAGGCACTGCTGGAGTGTG 181
 Db |||||
 QY 159 TCACAGATGACACGAGCCACCGAGGCTCCAGGGCCAGGAGGCACTGCTGGAGTGTG 218
 Db |||||
 QY 182 ACCACCTGAGAGTGGGGTCCAGGTGGAGGAGAGAGAGACTGGTCTGCTGCTCC 241
 Db |||||
 QY 219 ACCACCTGAGAGTGGGGTCCAGGTGGAGGAGAGAGAGACTGGTCTGCTGCTCC 278
 Db |||||
 QY 242 TCGTGGGCTCCCTCGCGGGCGCTTCGGCTCCTCTCTCTACGGCTACCACTCTCG 301
 Db |||||
 QY 279 TCGTGGGCTCCCTCGCGGGCGCTTCGGCTCCTCTCTCTCTACGGCTACCACTCTCG 338
 Db |||||
 QY 302 TGGTGAATGCCCCACCCGCTACATCAAGGCTTTTACAATGAGTCATGGGAAGAAGC 361
 Db |||||
 QY 339 TGGTGAATGCCCCACCCGCTCC----- 357
 Db |||||
 QY 362 ATGGACGTCCAATAGACCCAGACACTCTGACTCTGCTCTGGTCTGTGACTGTGTCCATAT 421
 Db |||||
 QY 358 ----- 357
 Db |||||
 QY 422 TCGCCATCGGTGGACTTGTGGGACGTTAATTGTGAAGATGATTGGAAGGTTCTTTGGGA 481
 Db |||||
 QY 358 ----- 357
 Db |||||
 QY 482 GGAAGCACATTTGCTGGCCAAATAATGGTTTCGAAATTTCTGCTGATTCGATGGCCT 541
 Db |||||
 QY 358 GGAAGCACATTTGCTGGCCAAATAATGGTTTCGAAATTTCTGCTGATTCGATGGCCT 417
 Db |||||
 QY 542 GCTCGCTCCAGGAGGAGCCCTTTGAAATGCTCATTTGGGAGCGCTTCATCATGGGCATAG 601
 Db |||||
 QY 418 GCTCGCTCCAGGAGGAGCCCTTTGAAATGCTCATCGTGGGACGCTTCATCATGGGCATAG 477
 Db |||||
 QY 602 ATGGAGGCGTCGCGCCTCAGTGTGCTCCCATGTACTCAGTGGAGATCTCACCCCAAGGAG 661
 Db |||||
 QY 478 ATGGAGGCGTCGCGCCTCAGTGTGCTCCCATGTACTCAGTGGAGATCTCACCCCAAGGAG 537
 Db |||||
 QY 662 TCGTGGCTCTCGGGCAGGAGTGCATGCTTTTATCTGCAATGCGCTGTTCACTGGGC 721
 Db |||||

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Loric, P., Legaspi, R.,
Maduro, Q. L., Mastello, C., Maskeri, B., Mastrian, S. D., McCloskey, J. C.,
McDowell, J., Pearson, R., Staniripoo, S., Thomas, P. J., Touchman, J. W.,
Turgeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 15 Row: b Column: 17.

FEATURES	source
----------	--------

```
1. .lib85
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="MGC:11313 IMAGE:3949549"
  /tissue_type="Ovary adenocarcinoma"
  /clone_lib="NIH MGC_9"
  /lab_host="DH10B-R"
  /vector="pOTB"
```

gene

```

/gene="SLC2A9"
/note="synonyms: GLUTX, Glut9"
/db_xref="LocusID:56606"
/db_xref="MIM:606142"
160.1695

```

SD

[illegible]

misc feature

```
241. .1614
/note="sugar_tr; Region: Sugar (and other) transporter"
/db xref="CDD:pfam00083"
```

ORIGIN

Query Match	61.7%	Score 1415.2;	DB 9;	Length 1885;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1450: Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

276

217 AGAAAGGACTGGTCTGGCTCGCTGGCTCCCTCGGGGCCCTTCGGCTCCTCC

017 33 4TTCGACCTCGCTCCTCGGCCTCCGTGGCGGCGGCTCGGTCCTCC 276

2217 AAATGGACTGGTCCTGCTCGCTCCCTCGAGGCCCTCCTCGGCCTCTTCCGCTCTCTCC

277 TTCTTCTACGGCTACACCTGTGGTGGTGAATGCCCCCAACCCCGTACATCAAGGCCCTT 336

[illegible]

TTCTCTACGGCTACAACCTGTGGTGGTGAATGCCCCCACCCCGTACATCAAGGCCTT 336

2 / 2

337 TACAAATGAGTCATGGGAAAGAAGGCATGGACGTCCAATAGACCCAGACACTCTGACTCTG 396

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

337 TACAATGAGTCATGGGAAGAAGGCATGGACGTCCAATAGACCCAGACACTCTGACTCTG 396

THE UNIVERSITY OF CHICAGO

397 CTCGGTCTGTGACTGTGTCCATATTCGCCATCGGTGGACTTGTGGGACGTTAATTGTG 456

397 CTCGGTCTGTGACTGTGTCCATATTCGCCATCGGTGGACTTGTGGGACATTAATTGTG 456

100

457 AAGATGATTGGAAAGGTTCTTGGGAGGAGCACACTTTGCTGGCCAAATAATGGGTTTGCA 516

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

457 AAGATGATTGGAAAGGTTCTTGGGAGGAAGCACACTTTGCTGGCCAATAATGGGTTTGCA 516

517 ATTTCTGCTGCA TTGCTGATGGCC TGTCTCGCTCCAGCAGGAGCCTTTGAAATGCTCATT 576

[illegible]

517 ATTCTGCTGCATTGCTGATGGCCTGCTCGCTCCAGCAGGAGCCTTTGAAATGCTCATC 576

Db	721	CGAGCTGCTGGGAAAGAGAGTACCTTGGCCATACCTGTTGGAGTGATGTTGGTCTCCCTG	780
QY	797	CGTGTCTCCAGCTGCTGAGCCCTCCCTTCTCCCGGACAGCCACAGCTACCTGCTCTGG	856
Db	781	CGTGTGTCAGTGTCTGAGCCCTCCCTTCTCCCGGACAGCCACAGCTACCTGCTCTGG	840
QY	857	AGAAGCACAAAGAGCAAGAGCTGTGAAGCCTTCCAAAGCTTCTTGGGTAAAGCAGACG	916
Db	841	AGAAGCACAAAGAGCAAGAGCTGTGAAGCCTTCCAAAGCTTCTTGGGTAAAGCAGACG	900
QY	917	TTTCCCMAGAGGTAGAGAGGTCTCTGCTGAGAGCCACGTCGAGAGGAGCATCCGCCCTGG	976
Db	901	TTTCCCAAGAGGTAGAGAGGTCTCTGCTGAGAGCCGCGTCGAGAGGAGCATCCGCCCTGG	960
QY	977	TGTCGCTGCTGAGCTGCTGAGAGCTCCCTACGTCGCTGGCAGGTGTCACCGTGATTG	1036
Db	961	TGTCGCTGCTGAGCTGCTGAGAGCTCCCTACGTCGCTGGCAGGTGTCACCGTGATTG	1020
QY	1037	TCACCATGGCCTGTACACAGCTCTGTGGCCTCAATGCAATTTGGTTCATACCAACAGCA	1096
Db	1021	TCACCATGGCCTGTACACAGCTCTGTGGCCTCAATGCAATTTGGTTCATACCAACAGCA	1080
QY	1097	TCCTTTGAAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACGTCACTGTGAGTACAGGG	1156
Db	1081	TCCTTTGAAAAGCTGGGATCCCTCTGGCAAAAGATCCCATACGTCACTGTGAGTACAGGG	1140
QY	1157	GCATCGAGACTTTGGCTGCGCTCTTCCTGCTGGTTGGTGCATTGAGCACCTGGGACGGAGAC	1216
Db	1141	GCATCGAGACTTTGGCTGCGCTCTTCCTGCTGGTTGGTGCATTGAGCACCTGGGACGGAGAC	1200
QY	1217	CCCTCCTCATTTGGTGGCTTTTGGGCTCATGGGCGCTCTTCTTTGGGACCCCTCACCATCACGC	1276
Db	1201	CCCTCCTCATTTGGTGGCTTTTGGGCTCATGGGCGCTCTTCTTTGGGACCCCTCACCATCACGC	1260
QY	1277	TGACCTTCGAGACACAGCCGCCCTCGGTGCCCTACCTGAGTATCGTGGGCAATTCGGCCCA	1336
Db	1261	TGACCTTCGAGACACAGCCGCCCTCGGTGCCCTACCTGAGTATCGTGGGCAATTCGGCCCA	1320
QY	1337	TCATCGCCTCTTCTGCAGTGGGCGAGGTGGCATCCCGTTCAATCTTTGACTGGTGAAGTCT	1396
Db	1321	TCATCGCCTCTTCTGCAGTGGGCGAGGTGGCATCCCGTTCAATCTTTGACTGGTGAAGTCT	1380
QY	1397	TCCAGCAATCTCAGCGCCCGGCTGCCCTCATCATTTGCAGGACACCGTCAACTGGCTCTCCA	1456
Db	1381	TCCAGCAATCTCAGCGCCCGGCTGCCCTCATCATTTGCAGGACACCGTCAACTGGCTCTCCA	1440
QY	1457	ACTTTGCTGTTGGGCTCCTCTTCCCATCATTCAGAAAAAGTCTGGACACCTACTGTGTTCC	1516
Db	1441	ACTTTGCTGTTGGGCTCCTCTTCCCATCATTCAGAAAAAGTCTGGACACCTACTGTGTTCC	1500
QY	1517	TAGTCTTTGCTACAAATTTGATCACAGGTGCTATCTACCTGATATTTTGTGCTGCTCGATGA	1576
Db	1501	TAGTCTTTGCTACAAATTTGATCACAGGTGCTATCTACCTGATATTTTGTGCTGCTCGATGA	1560
QY	1577	CCAAAAACAGAACCTATGCAGAAATCAGGCCAGGCAATTTTCCAAAAGGAAACAAAGCATACC	1636
Db	1561	CCAAAAACAGAACCTATGCAGAAATCAGGCCAGGCAATTTTCCAAAAGGAAACAAAGCATACC	1620
QY	1637	CACAGAGAGAAATCGACTCAGCTGTCACTGATGCTC	1675
b	1621	CACAGAGAGAAATCGACTCAGCTGTCACTGATGCTC	1659

RESULT 4
AF210317
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF210317
Homo sapiens
(SLC2A9) mRNA,
complete cds.
AF210317
AF210317.1
GI:9230650
Homo sapiens
(human)
Homo sapiens
ORGANISM

1863 bp
mRNA
linear
PRI 16-JUL-2000
Homo sapiens
facilitative glucose
transporter family member
GLUT9

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Phay, J.E., Hussain, H.B. and Moley, J.F.
Cloning and expression analysis of a novel member of the
facilitative glucose transporter family, SLC2A9 (GLUT9)
Genomics 66 (2), 217-220 (2000)
20318624
10860667
2 (bases 1 to 1863)
Phay, J.E. and Moley, J.F.
Direct Submission
Submitted (01-DEC-1999) Department of Surgery, Washington
University School of Medicine, 660 S. Euclid Box 8109, St. Louis,
MO 63110, USA

FEATURES source

```

source
1. .1863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p15.3-p16"
1. .1863
gene
/gene="SLC2A9"
55. .1677
CDS

```

gene

gene="SLC2A9"
55. 1677
CDS

CDs

/gene="SLC2A9"
 /codon_start=1
 /product="facilitative glucose transporter family member
 GLUT9"
 /protein_id="AAF85942.1"
 /db_xref="GI:9230851"
 /translation="MARQKNRSKELGVLTDDTSHARPPGQRALLECDHRSQ
 GRRRRKDWCSLLVASAGAFSSFLGYNLSVNVNAPTPYIKAFYNSWERRHQRP
 PQTLLWSTVYSIPAIIGLVGTLLVKMIGKVLGKTKLLANNGFATSAALLMACSI
 AGAFELVGRIMGIDGVALSVLPDMLYSISPKETKSGSQVTAIFICIGVTFG
 LGPELLGKSTWPTLYFGVIVVPAVLELLSPFLPDSPRYLLEKXNEARAVKQAL
 LQKADQVESEVLASRQVRSIRLVSVLELLRAIPYVRQWVTVITVMACQGLGL
 IWFYNSIPGKAGIPLAKIPYVTLSTGCIETLAAVFSGLVIEHLGRPLLLGGFLA
 LFFGLTITLTLDQDHPWPYLSIVGIIUAIASFCSGGGPIFLTGE7FQOOSORRP
 FIATGVNWSNFAYGLLFPPIQKSLDTPYCFELVATICITGAIYLVFVLPETKRT
 ETSQSSKKNKAYPEEKIEGAVTDKINGRP"

ORIGIN

Query Match 70.3%; Score 1646.4; DB 9; Length 1863;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1650; Conservative 0; Mismatches 6; Indels 0; Gaps 0

[illegible]

Db 1441 GTCAACTGGCTCTCCAACTTGTGTTGGGCTCCTCTTCCCATTCATTCAGAAAAGCTGTG 1500
QY 1501 GACACCTACTGTTTCTTCTAGTCTTCTGCTACAAATTTGTATATACAGGTGCTATCTACCTGTAT 1560
Db 1501 GACACCTACTGTTTCTTCTAGTCTTCTGCTACAAATTTGTATATACAGGTGCTATCTACCTGTAT 1560
QY 1561 TTGTGCTGCTGACACCAAAAACAGAACTTATCAGAAATTCAGGCAATTTTCCAAA 1620
Db 1561 TTGTGCTGCTGACACCAAAAACAGAACTTATCAGAAATTCAGGCAATTTTCCAAA 1620
QY 1621 AGGAACAAAGCATACCCACCAAGAGAGAAATTCAGCTCAGCTGCTAGTATGCTCTGCT 1680
Db 1621 AGGAACAAAGCATACCCACCAAGAGAGAAATTCAGCTCAGCTGCTAGTATGCTCTGCT 1680
QY 1681 TTTTCTCTTCTACATCTCCGAATACAGCTGGAATCAAGCTCCGCCACCAACCGCC 1740
Db 1681 TTTTCTCTTCTACATCTCCGAATACAGCTGGAATCAAGCTCCGCCACCAACCGCC 1740
QY 1741 ACCAAAAGACACCCATCTTAACGGTCAATGTTTCTTCTCACTTCACTTCACTTCACT 1800
Db 1741 ACCAAAAGACACCCATCTTAACGGTCAATGTTTCTTCTCACTTCACTTCACTTCACT 1800
QY 1801 TCCCTTATCTTCTTCTCTGAGAACCAACAGTATCATGATTCAGACAGAGCTTGGATTT 1860
Db 1801 TCCCTTATCTTCTTCTCTGAGAACCAACAGTATCATGATTCAGACAGAGCTTGGATTT 1860
QY 1861 TGGAGACATGGGTTTGAATTCAGTCAATTCATTTTATTCAGCAATATTTTAAACAT 1920
Db 1861 TGGAGACATGGGTTTGAATTCAGTCAATTCATTTTATTCAGCAATATTTTAAACAT 1920
QY 1921 ACTGACATGCTCCATATGTTTATCCACTGGTTATACAAATGGGAGGAGAGAG 1980
Db 1921 ACTGACATGCTCCATATGTTTATCCACTGGTTATACAAATGGGAGGAGAGAG 1980
QY 1981 AGAGAGAGAGAGAGATGCTTATCTAAAGCTTGAAGTCTAGGCTGTGCAGGCTGCTC 2040
Db 1981 AGAGAGAGAGAGAGATGCTTATCTAAAGCTTGAAGTCTAGGCTGTGCAGGCTGCTC 2040
QY 2041 AGCCTGTATCTCCAGACATCTTGGAGGCGAGTGGTGGATCTGAGTCAAGAGATT 2100
Db 2041 AGCCTGTATCTCCAGACATCTTGGAGGCGAGTGGTGGATCTGAGTCAAGAGATT 2100
QY 2101 GAGACATCTCTGGCTTAAATGCTGAACTCCCTCTCTACTTAAATAACAAAATTAGCT 2160
Db 2101 GAGACATCTCTGGCTTAAATGCTGAACTCCCTCTCTACTTAAATAACAAAATTAGCT 2160
QY 2161 GAGCATGCTGCGGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGAGGAGAGATGCG 2220
Db 2161 GAGCATGCTGCGGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGAGGAGAGATGCG 2220
QY 2221 GTGAACCCAGAGGCGGAGCTTGCAGTGAAGCGGAGATCACACCACTCCAGCTGG 2280
Db 2221 GTGAACCCAGAGGCGGAGCTTGCAGTGAAGCGGAGATCACACCACTCCAGCTGG 2280
QY 2281 GTGACAGAGCCAGACTCCGCTCTCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 2281 GTGACAGAGCCAGACTCCGCTCTCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
QY 2341 CGC 2343
Db 2341 CGC 2343

RESULT 3
AX709981
LOCUS AX709981 2032 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 52 from Patent WO0192304.
ACCESSION AX709981
VERSION AX709981.1 GI:29786598
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
Bruns, C.M., Seilhamer, J.J., Tribouley, C.M., Azimzai, Y., Ding, L.,
Hafalia, A., Khan, F.A., Yao, M.G., Yue, H., Au-Young, J., Baughn, M.R.,
Borowsky, M.L., Buford, N., Elliot, V.S., Gandhi, A.R., Green, B.D.,
Kearney, L., Lai, P., Lee, E.A., Lu, D.A.M., Lu, Y., Tang, Y.T., Yang, J.,
Griffin, J.A., Hillman, J.B., Nguyen, D.B., Policky, J.L., Ramkumar, J.,
Raumann, B.E., Sanjanwalla, M.S., Thornton, M. and Wallia, N.K.
Transposers and ion channels
Patent: WO 0192304-A 52 06-DEC-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1. .2032
/organism="Homo sapiens"
/mol type="unassigned DNA"
/db xref="taxon:9606"
/note="Incyte ID No: 7476643CB1"

ORIGIN

Query Match 70.5%; Score 1651; DB 6; Length 2032;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 17 GCCTTGGCAGAGTCTGGGCTCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 76
Db 1 GCCTTGGCAGAGTCTGGGCTCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 60
QY 77 CAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCCCTCTCACAGATGACACCA 136
Db 61 CAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCCCTCTCACAGATGACACCA 120
QY 137 GCCACGCCGGGCTCCAGGGCCAGGAGGAGCACTCTGGAGTGTGACCACTGAGGAGTG 196
Db 121 GCCACGCCGGGCTCCAGGGCCAGGAGGAGCACTCTGGAGTGTGACCACTGAGGAGTG 180
QY 197 GGGTGCAGAGTGGAGGAG 256
Db 181 GGGTGCAGAGTGGAGGAG 240
QY 257 CGGGCGCTTCCGCT 316
Db 241 CGGGCGCTTCCGCT 300
QY 317 CCCGTGATCTAAGGCTTTTACATGATGATCATGGGAAAGAGGATGAGCTCAATAG 376
Db 301 CCCGTGATCTAAGGCTTTTACATGATGATCATGGGAAAGAGGATGAGCTCAATAG 360
QY 377 ACCGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 436
Db 361 ACCGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 420
QY 437 TTGTGGGAGCTTAATTTGAGATGATGGAAGGTTCTTGGAGGAGGAGCACTTTGC 496
Db 421 TTGTGGGAGCTTAATTTGAGATGATGGAAGGTTCTTGGAGGAGGAGCACTTTGC 480
QY 497 TGGCCCAATAGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
Db 481 TGGCCCAATAGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 557 GAGCCTTTGAAATGCTCAATTTGTGGAGCTTTTATCATGGGATAGATGAGGCGCTGCCC 616
Db 541 GAGCCTTTGAAATGCTCAATTTGTGGAGCTTTTATCATGGGATAGATGAGGCGCTGCCC 600
QY 617 TCAAGTGTCTCCCATGTACCTCAGTGAATCTACCCAGGAGATCGTGGCTCTCTGG 676
Db 601 TCAAGTGTCTCCCATGTACCTCAGTGAATCTACCCAGGAGATCGTGGCTCTCTGG 660
QY 677 GGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACCTGGGAGGATCTTGGGCGCTGC 736
Db 661 GGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACCTGGGAGGATCTTGGGCGCTGC 720
QY 737 CCGAGCTGTGGGAAAGAGAGATACCTGCCATACCTGTTTGGAGTGAATTTGTGCTGCTG 796

[illegible]

361	Db		CATGACGTC	CAATAGACCCAGACACTCTGCTCTGCTGTGACTGTGTCCATA	420
421	QY		TTTCGCCATCG	CGTGACCTTGTGGGAGCGTTAAATTGTGTGAAGATGATTGGAAAGGTTCTTGGG	480
421	Db		TTTCGCCATCG	CGTGACCTTGTGGGAGCGTTAAATTGTGTGAAGATGATTGGAAAGGTTCTTGGG	480
481	QY		AGGAAGACAC	ATTGTGTGCGCCAAATAATATGGGTTTGCAATTTCTGTGTCATGTCTGATGSCC	540
481	Db		AGGAAGACAC	ATTGTGTGCGCCAAATAATGGTTTGCAATTTCTGTGTCATGTCTGATGSCC	540
541	QY		TGCTCGCTCC	AGGACAGAGAGCGTTTGAATGCTCATGTGGAGACGCTTCATCATGGGCATA	600
541	Db		TGCTCGCTCC	AGGACAGAGAGCGTTTGAATGCTCATGTGGAGACGCTTCATCATGGGCATA	600
601	QY		GATGAGAGCG	GTGCGCTCAGTGTGTCCCCATGTATACCTCAGTGAGATCTCACCCCAAGGAG	660
601	Db		GATGAGAGCG	GTGCGCTCAGTGTGTCCCCATGTATACCTCAGTGAGATCTCACCCCAAGGAG	660
661	QY		ATCCGTGGCT	CTCTGGGGCAGGTGACTGCATCTTTATCTGCATTTGGGTGTTCACGTGGG	720
661	Db		ATCCGTGGCT	CTCTGGGGCAGGTGACTGCATCTTTATCTGCATTTGGGTGTTCACGTGGG	720
721	QY		CAGCTTCTG	GGCCCTGCCGAGCTGTGCGAAAGAGAGTACTGTGCCCATACCTGTTTGG	780
721	Db		CAGCTTCTG	GGCCCTGCCGAGCTGTGCGAAAGAGAGTACTGTGCCCATACCTGTTTGG	780
781	QY		GTGATTGTG	TCCTCGCTTGTCCAGCTGTGTAGCTTCCCTTTCTCCCGGACAGCCCA	840
781	Db		GTGATTGTG	TCCTCGCTTGTCCAGCTGTGTAGCTTCCCTTTCTCCCGGACAGCCCA	840
841	QY		CGCTACCTG	CTCTTTGGAGAGACACACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTTT	900
841	Db		CGCTACCTG	CTCTTTGGAGAGACACACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTTT	900
901	QY		TTGGTTAAAG	CAGACGTTTCCCAAGAGGTAGAGGAGTCTGTGGCTGAGAGCCACGTGCAG	960
901	Db		TTGGTTAAAG	CAGACGTTTCCCAAGAGGTAGAGGAGTCTGTGGCTGAGAGCCACGTGCAG	960
961	QY		AGGAGATCC	CGCTGCTGTGCGTCTGGAGCTGTGAGAGCTCCCTACGTCCCTGSCAG	1020
961	Db		AGGAGATCC	CGCTGCTGTGCGTCTGGAGCTGTGAGAGCTCCCTACGTCCCTGSCAG	1020
1021	QY		GTGCTCAC	CGTGATTGTCAACATGGCCCTGCTACCAAGCTGTGTGGCTCAAATGCAATTTGG	1080
1021	Db		GTGCTCAC	CGTGATTGTCAACATGGCCCTGCTACCAAGCTGTGTGGCTCAAATGCAATTTGG	1080
1081	QY		TTCTATACCA	AGACATCTTTGGAAAGAGTGGTGGATCCCTCCGGCAAGATCCCATACGTC	1140
1081	Db		TTCTATACCA	AGACATCTTTGGAAAGAGTGGTGGATCCCTCCGGCAAGATCCCATACGTC	1140
1141	QY		ACCTTGAGTA	CACGGGCGCATCGACACTTGGCTGCGGTCCTCTCTGGTTTGGTCAATTGAG	1200
1141	Db		ACCTTGAGTA	CACGGGCGCATCGACACTTGGCTGCGGTCCTCTCTGGTTTGGTCAATTGAG	1200
1201	QY		CACCTGGGA	CGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCCCTCTCTTTGGG	1260
1201	Db		CACCTGGGA	CGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCCCTCTCTTTGGG	1260
1261	QY		ACCTTCACCA	TCACTCAGCTGACCCCTGACGAGACACAGCCCTCGGGTCCCTTACCTGAGTATC	1320
1261	Db		ACCTTCACCA	TCACTCAGCTGACCCCTGACGAGACACAGCCCTCGGGTCCCTTACCTGAGTATC	1320
1321	QY		GTGGGCATT	CTGSCCATATCGCTCTTTCTGTAGTGGGCCAGGTGGCATCCCGTTTCATC	1380
1321	Db		GTGGGCATT	CTGSCCATATCGCTCTTTCTGTAGTGGGCCAGGTGGCATCCCGTTTCATC	1380
1381	QY		TTGACTGGT	GAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCATCTATGTCAGGACCC	1440
1381	Db		TTGACTGGT	GAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCATCTATGTCAGGACCC	1440
1441	QY		GTCAACTGG	CTCTCCAACCTTTTCTGTGGGCTCCTCTTCCCATTTCAATTCAGAAAAGCTGTG	1500

961	Qy	AGGAGCATCCCGCTGGTCTCGCTGCTGAGCTGCTGAGAGCTCCCTTACGTCCTCGCTGGCAG	1021
961	Db	AGGAGCATCCCGCTGGTGTCCGTGTGCTGAGCTGCTGAGAGCTCCCTTACGTCCTCGCTGGCAG	1020
1021	Qy	GTGGTCAACGCTGATTTGTCAACATGGCCCTGTACCAAGCTCTGTGGCCCTCAATGCAATTTGG	1080
1021	Db	GTGGTCAACGCTGATTTGTCAACATGGCCCTGTACCAAGCTCTGTGGCCCTCAATGCAATTTGG	1080
1081	Qy	TTCTATACCAACAGCATCTTTTGAAAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACGTC	1140
1081	Db	TTCTATACCAACAGCATCTTTTGAAAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACGTC	1140
1141	Qy	ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGGCTCTCTCTGTTTGGTTCATTGAG	1200
1141	Db	ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGCGGCTCTCTCTGTTTGGTTCATTGAG	1200
1201	Qy	CACCTGGAGCGAGACCCCTCCTCATTTGTTGGCTTTTGGCTCATGGGCCCTCTCTTTGGG	1260
1201	Db	CACCTGGAGCGAGACCCCTCCTCATTTGTTGGCTTTTGGCTCATGGGCCCTCTCTTTGGG	1260
1261	Qy	AGCCTCACCATCACGCTGACCCCTGAGGACCAACGCCCCCTGGGTCCCCTACCTGAGTATC	1320
1261	Db	AGCCTCACCATCACGCTGACCCCTGAGGACCAACGCCCCCTGGGTCCCCTACCTGAGTATC	1320
1321	Qy	GTGGGCATCTGGGCCCATCATGCGCTCTTTCTGAGTGGGCCAGGTGGCATCCCGTTCATC	1380
1321	Db	GTGGGCATCTGGGCCCATCATGCGCTCTTTCTGAGTGGGCCAGGTGGCATCCCGTTCATC	1380
1381	Qy	TTGACTGTGAGTTCCTTCCAGCAATCTCAGCGGCGGCTGCCCTCATCATTTGAGGACACC	1440
1381	Db	TTGACTGTGAGTTCCTTCCAGCAATCTCAGCGGCGGCTGCCCTTCATCATTTGAGGACACC	1440
1441	Qy	GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCCTCTTCCCATTCAATCAGAAAACTCTG	1500
1441	Db	GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCCTCTTCCCATTCAATCAGAAAACTCTG	1500
1501	Qy	GACACCTACTGTTTCCCTAGTCTTTGCTACAAATTTGTATCACAGGTGCTATCTACCTGTAT	1560
1501	Db	GACACCTACTGTTTCCCTAGTCTTTGCTACAAATTTGTATCACAGGTGCTATCTACCTGTAT	1560
1561	Qy	TTTGTGCTGCTGAGACCAAAACAGAACCTATGCGAAATCAGCCAGGCAATTTTCCAAA	1620
1561	Db	TTTGTGCTGCTGAGACCAAAACAGAACCTATGCGAAATCAGCCAGGCAATTTTCCAAA	1620
1621	Qy	AGGAACAAGCATACCCACCAAGAGAAAATCGACTCAGCTGTCTACTGATGCTCCTGCT	1680
1621	Db	AGGAACAAGCATACCCACCAAGAGAAAATCGACTCAGCTGTCTACTGATGCTCCTGCT	1680
1681	Qy	TCCTCTCCTTTCACTACTCCGAATACAGCCTGGATTCAAAGCTGCCGCCACCAACCCGCC	1740
1681	Db	TCCTCTCCTTTCACTACTCCGAATACAGCCTGGATTCAAAGCTGCCGCCACCAACCCGCC	1740
1741	Qy	ACCAAAAAAGAACACCCATTGTAACCGGTCAATGCGTATTTCTCTAACTTGGAATGACCT	1800
1741	Db	ACCAAAAAAGAACACCCATTGTAACCGGTCAATGCGTATTTCTCTAACTTGGAATGACCT	1800
1801	Qy	TCCCTTATCTCTTCTCGGAGAACACCAAGTCATGATCTCAGACAAGAGCTTGGAATTT	1860
1801	Db	TCCCTTATCTCTTCTCGGAGAACACCAAGTCATGATCTCAGACAAGAGCTTGGAATTT	1860
1861	Qy	TGGAGACATGGGTTTGAAATTCAGATCATTTCTTTTATTCAGCAAAATTTAAACAAGT	1920
1861	Db	TGGAGACATGGGTTTGAAATTCAGATCATTTCTTTTATTCAGCAAAATTTAAACAAGT	1920
1921	Qy	ACTGACATGTCCCATATGTTGTTTTTACCACTGGTTTATACAAATGGGAGGAGAGAGAG	1980
1921	Db	ACTGACATGTCCCATATGTTGTTTTTACCACTGGTTTATACAAATGGGAGGAGAGAGAG	1980
1981	Qy	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGACCGTGGCTC	2040
1981	Db	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGACCGTGGCTC	2040
2041	Qy	AGCCCTCTATCTCCACGACTTTTGGAGGGCGGCTGGTGGTGGATCTGATCTGAGCTCAGGAGATT	2100

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 7, 2004, 21:03:37 ; Search time 9011 Seconds
(without alignments)
11269.862 Million cell updates/sec
Title: US-09-981-947B-1
Perfect score: 2343
Sequence: 1 tgcaccacgcgtccggcct.....aaaaaaaaagggcgccgcg 2343

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues 6940544
Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2343	100.0	2343	6	AR135669 Sequence
2	2343	100.0	2343	6	AR184379 Sequence
3	1651	70.5	2032	6	AX709981 Sequence
4	1646.4	70.3	1863	9	AF210317 Homo sapi
5	1445.2	61.7	1885	9	BC018897 Homo sapi
6	1317.8	56.2	2842	6	AX339083 Sequence
7	1141	48.7	2216	6	AX179745 Sequence
8	1089.2	46.5	2022	10	AF469480 Mus muscu
9	636.6	27.2	160402	2	AC067775 Homo sapi
10	636.6	27.2	165221	9	AC098976 Homo sapi
11	616.2	26.3	3682	9	HSU21164 Human dopam
12	592.2	25.3	597	9	AF421859 Homo sapi
13	523.6	22.3	1296	10	AF490463 Mus muscu
14	515.8	22.0	3243	10	BC006076 Mus muscu
15	427.6	18.3	1539	6	AX753091 Sequence
16	427	18.2	153113	9	AC006453 Homo sapi
17	424.8	18.1	1714	10	RATGT5 D28562 Rattus norv
18	424.4	18.1	1539	6	AX686979 Sequence
19	424.4	18.1	2080	10	BC023500 Mus muscu
20	423.4	18.1	2137	10	RATGLUTV L05195 Rattus norv
21	423.4	18.1	138007	2	AC141282 Homo sapi
22	423.4	18.1	173505	2	AC141297 Homo sapi
23	423.4	18.1	213386	2	AC139261 Homo sapi
24	422	18.0	105643	9	AC005629 Homo sapi
25	421.6	18.0	2169	10	RATGLUTS D13871 Rattus norv
26	421.2	18.0	2069	10	AF161071 Mus muscu
27	421	18.0	168754	2	AC027487 Homo sapi
28	419.4	17.9	158052	9	AC027612 Homo sapi
29	419.4	17.9	159575	2	AC141407 Homo sapi
30	419.4	17.9	165225	9	AC116050 Homo sapi
31	419.4	17.9	195985	2	AC141251 Homo sapi
32	419.4	17.9	197109	2	AC141252 Homo sapi
33	416.4	17.8	2213	9	BC001692 Homo sapi
34	416.4	17.8	2213	9	BC001820 Homo sapi
35	416.4	17.8	2218	6	AX332575 Sequence
36	416.4	17.8	2218	6	AX333214 Sequence
37	416.4	17.8	2218	9	HUMGLUT5 M55531 Human glucoc
38	416	17.8	155956	2	AC144898 Homo sapi
39	416	17.8	193913	2	AC145614 Homo sapi
40	414	17.7	57598	9	AL513526 Human DNA
41	410	17.5	7434	6	AX348519 Sequence
42	407.8	17.4	7434	6	AX348520 Sequence
43	402	17.2	2166	4	ECA555215 Equus cab
44	390.8	16.7	2147	4	OAR315928 Ovis arie
45	376	16.0	1607	6	AX556510 Sequence

ALIGNMENTS

RESULT 1
AR135669
LOCUS AR135669 2343 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6136547.
ACCESSION AR135669
VERSION AR135669.1 GI:14476341
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2343)
Tartaglia, L.A. and Weng, X.
AUTHORS Nucleic acid molecules encoding glutx and uses thereof
TITLE Patent: US 6136547-A 1 24-OCT-2000;
JOURNAL Location/Qualifiers
FEATURES

Db	394	TACATGAGTCATGGGAAAGAGCATGGACGTCCATAGACCCAGACACTCTGACTCTG	453
Qy	397	CTCTGGTCGTGACTGTGTCCATATTCCGCAATCGGTGGACTTTGTGGGACGTTAATTGTG	456
Db	454	CTCTGGTCGTGACTGTGTCCATATTCCGCAATCGGTGGACTTTGTGGGACATTAATTGTG	513
Qy	457	AAGATGATTGGAAAGGTTCTTGGGAGGAAAGCACACTTTGCTGGCCCAATAATGGGTTTGCA	516
Db	514	AAGATGATTGGAAAGGTTCTTGGGAGGAAAGCACACTTTGCTGGCCCAATAATGGGTTTGCA	573
Qy	517	ATTTCTGCTGCATTGTGATGGCTCTCGCTCCAGGACAGGAGCCCTTTGAAATGCTCATT	576
Db	574	ATTTCTGCTGCATTGTGATGGCTCTCGCTCCAGGACAGGAGCCCTTTGAAATGCTCATT	632
Qy	577	GTGGGACGTTTCATATGGGCATAGATGGAGGCGTCCCTCAGTGTGCTCCCATGTAC	636
Db	633	GTGGGACGTTTCATATGGGCATAGATGGAGGCGTCCCTCAGTGTGCTCCCATGTAC	691
Qy	637	CTCAGTGAGATCTCACCCAGGAGATCCGTTGGCTCTCTGGGGCAGGTGACTGCCATCTTT	696
Db	692	CTCAGTGAGATCTCACCCAGGAGATCCGTTGGCTCTCTGGGGCAGGTGACTGCCATCTTT	750
Qy	697	ATCTGCATTGGCG	709
Db	751	ATCTGCATTGGCG	763

Search completed: April 8, 2004, 16:08:51
Job time : 5995 secs